МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования
"Юго-Западный государственный университет"
(ЮЗГУ)

УТВЕРЖДАЮ
Проректор по учебной работе
_____________ О.Г. Локтионова
"_4_" _04_ ____________ 2022 г.

МОДЕЛИРОВАНИЕ И ПРОЕКТИРОВАНИЕ ЭЛЕКТРОЭНЕРГЕТИЧЕСКИХ ПРОЦЕССОВ

Методические указания у практическим занятиям по дисциплине "Моделирование и проектирование электроэнергетических процессов" для студентов всех форм обучения направления подготовки 13.04.02 Электроэнергетика и электротехника

Курск 2022

УДК 621.311.

Составитель А.В. Филонович, И.В. Ворначева

Рецензент:

Кандидат технических наук, доцент кафедры «Электроснабжение»

О.М. Ларин

Моделирование и проектирование электроэнергетических процессов: методические указания по выполнению практических работ / Юго-Зап. гос. ун-т; сост.: А.В. Филонович, И.В. Ворначева. Курск, 2022. 52 с., Библиогр.:

Содержат сведения по выполнению практических работ по дисциплине «Моделирование и проектирование электроэнергетических процессов».

Методические указания соответствуют требованиям программы, утвержденной учебно-методическим объединением для направления подготовки 13.04.02 Электроэнергетика и электротехника

Предназначены для студентов всех форм обучения.

Текст печатается в авторской редакции

Подписано в печать . Формат 60x84 1/16.

Усл.печ.л. . Уч.–изд.л . Тираж 100 экз. Заказ . Бесплатно.

Юго-Западный государственный университет.
305040, г.Курск, ул.50 лет Октября, 94

ПРАКТИЧЕСКОЕ ЗАНЯТИЕ №1

РЕШЕНИЕ ЗАДАЧ ОПТИМИЗАЦИИ С ПОМОЩЬЮ НАДСТРОЙКИ «ПОИСК РЕШЕНИЯ В СРЕДЕ MICROSOFT EXCEL

В экономике и управлении оптимизационные задачи возникают в связи с многочисленностью возможных вариантов функционирования конкретного экономического объекта, когда возникает ситуация выбора варианта, наилучшего по некоторому правилу, критерию, характеризуемому соответствующей целевой функцией (например, минимум затрат, максимум продукции).

***Задача №1 «*Задача оптимального использования ресурсов»**

Фабрика имеет в своем распоряжении определенное количество ресурсов: рабочую силу, сырье и оборудование, производственные площади и т.п. Допустим, например, ресурсы трех видов: рабочая сила, сырье и оборудование – имеются в количестве соответственно 80(чел./дней), 480(кг) и 130 (станко/час). Фабрика может выпускать ковры четырех видов. Информация о количестве единиц каждого ресурса, необходимых для производства одного ковра каждого вида и доходах, получаемых предприятием от единицы каждого вида товаров, приведена в таблице.

Ресурсы	Нормы ресурсов на единицу изделия				Наличие ресурсов
	Ковер «Лужайка»	Ковер «Силуэт»	Ковер «Детский»	Ковер «Дымка»	
Труд	7	2	2	6	80
Сырье	5	8	4	3	480
Оборудование	2	4	1	8	130
Цена (тыс.руб.)	3	4	3	1	

Требуется найти такой план выпуска продукции, при котором будет максимальной общая стоимость продукции.

Формализации задачи

Обозначим через X_1, X_2, X_3, X_4 количество ковров каждого вида. Тогда прибыль, которую может получить предприятие, составит:

$F(x)=3\ X_1+4\ X_2+3\ X_3+\ X_4$

Необходимо найти такие X_1, X_2, X_3, X_4 при которых F(x) достигнет максимума.

Выражение, которое необходимо максимизировать или минимизировать называют целевой функцией.

Установка границ оптимизируемой системы:

Предприятие имеет ограниченные ресурсы, которые невозможно превысить.

$7\ X_1+2\ X_2+2\ X_3+\ 6X_4 \leq 80$ – ограничение на трудовые ресурсы

$5\ X_1+8\ X_2+4\ X_3+\ 3X_4 \leq 480$ – ограничение на сырьевые ресурсы

$2X_1+4\ X_2+\ X_3+\ 8X_4 \leq 130$ – ограничение на ресурсы оборудования

Найденное решение должно быть неотрицательным.

Решение:

Допустим, что предприятие выпускает по одному ковру каждого вида.

1.На рабочем листе Microsoft Excel создайте следующую таблицу

	A	B	C	D	E	F	G
		Ковер «Лужайка»	Ковер «Силуэт»	Ковер «Детский»	Ковер «Дымка»		
	кол-во	1	1	1	1		
	цена	3	4	3	1		
	Целевая функция						
	ресурсы					расчетное кол-во ресурсов	Наличие ресурсов
	труд	7	2	2	6		80
	сырье	5	8	4	3		480
	оборудование	2	4	1	8		130

2. В ячейку B4 введем целевую функцию: =B3*B2+C3*C2+D3*D2+E3*E2

Мы узнаем, какой доход получит предприятие, если произведет по одному ковру каждого вида.

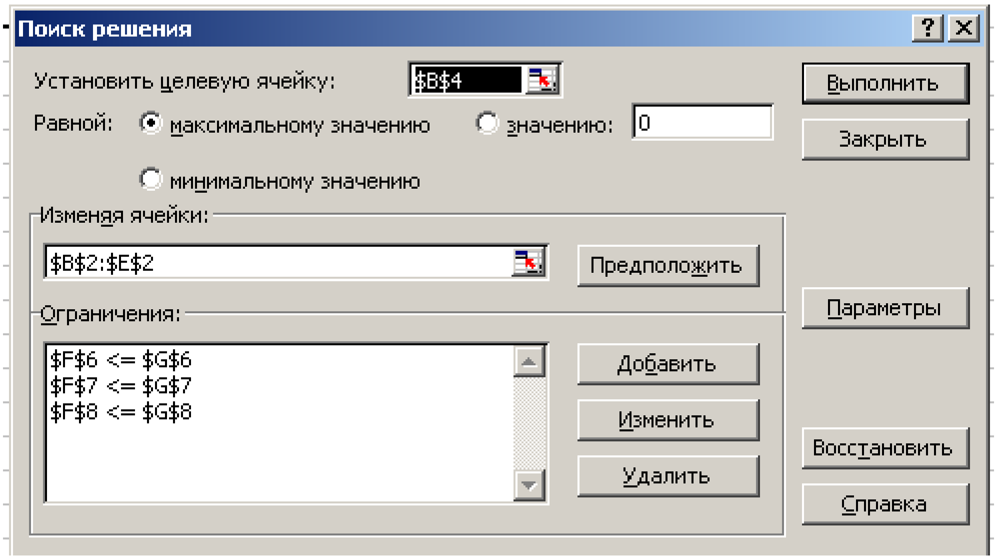
3. Рассчитаем количество ресурсов, которые будут затрачены, если предприятие произведет по одной единице, каждой продукции. Запишем в F6 затраты труда, определив их формулой: =B6*$B$2+C6*$C$2+D6*$D$2+E6*$E$2

В F7 затраты сырья, определив их формулой: =B7*$B$2+C7*$C$2+D7*$D$2+E7*$E$2

В F8 использование ресурсов оборудования, определив их формулой: =B8*$B$2+C8*$C$2+D8*$D$2+E8*$E$2

3. Выделите ячейку с формулой целевой функции B4 и выполните команду Данные/Поиск решения.

В диалоговом окне ***Поиск Решения*** разместим условие задачи

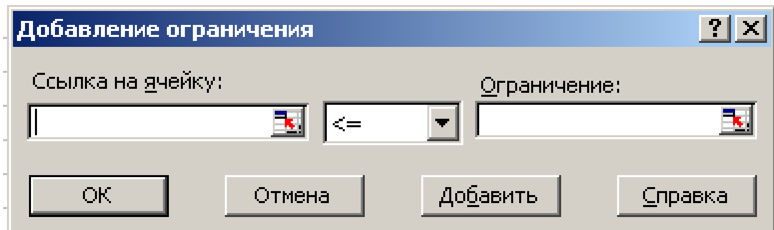


В поле «установить ячейку» введем адрес $B$4

В группе переключателей выберем для целевой ячейки «максимальному значению».

В поле «изменяя ячейки» с помощью мыши или клавиатуры ведем адреса изменяемых ячеек.

Щелкнув по кнопке Добавить, вызовем окно диалога «Добавление ограничения»



Установим первое ограничение, для чего введем ссылку на ячейку$F$6, выберем оператор <=, укажем адрес $G$6 в ограничение. Далее щелкнем по кнопке Добавить и введем следующее ограничение. После ввода всех ограничений нажмем ОК.

Щелкнем кнопку Параметры и для настройки итерационного цикла, а также ускорения поиска включим параметры: *линейная модель, неотрицательные значения.* Остальные параметры оставим по умолчанию.

Нажмем кнопку Выполнить – и программа выведет результаты решения:

Максимальная прибыль равна 150. При это будет необходимо произвести 30 ковров «Силуэт» и 10 ковров «Детский». Кроме этого на экран будет выведено диалоговое окно *Результаты решения*, в котором EXCEL сообщит, что все условия оптимальности выполнены и решение найдено. В этом же окне будет предложено сохранить найденное решение.

Задачи для самостоятельного решения

Задача №1

На заводе выпускают изделия четырех типов. От реализации 1ед. каждого изделия получает прибыль соответственно 2, 1, 3, 5 д.е. на изготовление изделий расходуются ресурсы трех типов: энергия, материалы, труд. Данные о технологическом процессе приведены в таблице

Ресурсы	Затраты ресурсов на единицу изделия				Запасы ресурсов, ед.
	I	II	III	IV	
Энергия	2	3	1	2	30
Материалы	4	2	1	2	40
Труд	1	2	3	1	25
Цена (тыс.руб.)	2	1	3	5	

Спланируйте производство изделий так, чтобы прибыль от реализации была наибольшей.

Задача №2

Предприятие должно выпускать два вида продукции - А и В, используя при этом последовательно четыре станка. Данные о технологическом процессе указаны в следующей таблице:

Станок	Трудоемкость на 1 ед. продукции		Фонд времени, час
	A	B	
1	3	3	15
2	2	6	18
3	4	0	16
4	1	2	8
Прибыль на 1 ед. продукции	2	3	

Спланируйте производство изделий так, чтобы прибыль от реализации была наибольшей.

Задача №3

Из трех продуктов I, II, III составляется смесь. В состав смеси должно входить не менее 6 ед. химического вещества А, 8 ед. – вещества В и не менее 12 ед. вещества С. Структура смеси приведена в следующей таблице:

Продукт	Содержание химического вещества в 1 ед. продукции			Стоимость 1 ед. продукции
	A	B	C	
I	2	1	3	2
II	1	2	4	3
III	3	1.5	2	2.5

Составьте наиболее дешевую смесь.

Задача №4

При откорме каждое животное должно получить не менее 9 ед. белков, 8 ед. углеводов и 11 ед протеина. Для составления рациона используют два вида корма, представленных в следующей таблице:

Питательные вещества	Количество единиц питательных веществ на !кг	
	Корма 1	Корма 2
Белки	3	1
Углеводы	1	2
Протеина	1	6

Стоимость 1 кг корма первого вида – 4 д.е., второго – 6 д.е.

Составьте дневной рацион питания, имеющий минимальную стоимость.

Транспортная задача.

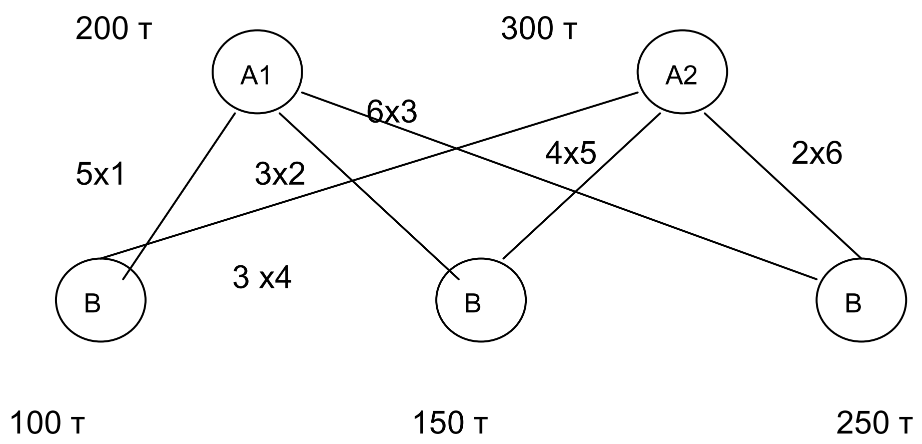
Имеются два склада готовой продукции: A_1 и A_2 с запасами однородного груза 200 и 300т. Этот груз необходимо доставить трем потребителям: B_1, B_2 и B_3 в количестве 100, 150, 250т соответственно. Стоимость перевозки 1т груза со склада A_1 потребителям B_1, B_2 и B_3 равна 5,3,6 д.е., а из склада A_2 тем же потребителям – 3,4,2, д.е соответственно. Составьте план перевозок, минимизирующий суммарные транспортные расходы.

Решение:

Обозначим за X1, X2, X3 – количество тонн груза, которое надо перевести со склада A1 каждому потребителю.

Обозначим за X4, X5, X6 – количество тонн груза, которое надо перевести со склада A2 тем же потребителям. Тогда транспортные расходы можно выразить функцией:

F(x)=5x1+3x2+6x3+3x4+4x5+2x6 $\rightarrow$ min



Ограничения:

X1+X2+X3<=200

X4+X5+X6<=300

X1+X4>=100

X2+X5>=150

X5+X6>=250

Контрольные задачи

Вариант 1.

1. Имеются два элеватора, в которых сосредоточено соответственно 4000 и 1500 тонн зерна. Зерно необходимо перевезти трем хлебозаводам в количестве 1000, 2000 и 1500 тонн каждому. Расстояние от элеватора до хлебозаводов указано в следующей таблице

Элеваторы	Хлебозаводы		
	1	2	3
1	20	30	20
2	50	40	40

Затраты на перевозку 1 тонны зерна на 1 км составляют 25 д.е. Спланируйте перевозки зерна из условия минимизации транспортных расходов.

2. Цех выпускает трансформаторы двух видов. Для изготовления трансформаторов обоих видов используются железо и проволока. Общий запас железа – 3 т, проволоки – 18 т. На один трансформатор первого вида расходуется 5 кг железа и 3 кг проволоки, а на один трансформатор второго вида расходуется 3 кг железа и 2 кг проволоки. За каждый реализованный трансформатор первого вида завод получает прибыль 3 д.е., второго – 4 д.е. Составьте план выпуска трансформаторов, обеспечивающий заводу максимальную прибыль.

Вариант 2.

1. Компания производит полки для ванных комнат двух размеров - A и B. Агенты по продаже считают, что в неделю на рынке может быть реализовано до 550 полок. Для каждой полки типа A требуется 2 м2 материала, а для полки типа B - 3 м2 материала. Компания может получить до 1200 м2 материала в неделю. Для изготовления одной полки типа A требуется 12 мин машинного времени, а для изготовления одной полки типа B - 30 мин; машины можно использовать 160 ч в неделю. Если прибыль от продажи полок типа A составляет 3 дол, а от полок типа B - 4 дол., то сколько полок каждого типа следует выпускать в неделю?

2.В некоторой местности в двух пунктах A и B имеется потребность в дополнительном транспорте. В пункте А требуется 5 дополнительных автобусов, а в пункте В - 7. Известно, что 3, 4 и 5 автобусов могут быть получены соответственно из гаражей Г1, Г2, и Г3. Как следует распределить эти автобусы между пунктами А и В, чтобы минимизировать их суммарный пробег? Расстояния от гаражей до пунктов А и В приведены в таблице.

Гараж	Расстояние до пунктов	
	А	В
Г1	3	4
Г2	1	3
Г3	4	2

Контрольные вопросы

1. Назначение оптимизационных задач.
2. Методика решения оптимизационных задач в Microsoft Excel.
3. Что такое формализация задачи?
4. Установка границ оптимизации системы.

ПРАКТИЧЕСКОЕ ЗАНЯТИЕ №2
ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ УПРАВЛЕНИЯ ЗАПАСАМИ МЕТОДОМ МОНТЕ-КАРЛО

Постановка задачи. Корпорация занимается производством некоторых изделий. Для их производства необходимы детали (аккумуляторы), которые закупаются у поставщика. На основе прошлого опыта специалисты оценили, что спрос за 100 недель колеблется от 670 до 740 штук. Частота спроса на аккумуляторы показана в таблице 3.1.

Таблица 3.1. Частота спроса на аккумуляторы

Спрос в неделю	**Частота**
670	3
675	7
680	5
685	10
690	13
695	8
700	5
705	2
710	13
715	10
720	9
725	3
730	7
735	2
740	3

Начальный запас деталей составляет 1800 шт., причем администрация компании приняла решение о подачах заказов на партии деталей размером в 2500 шт. каждый раз, когда их запас опускается ниже уровня в 1300 шт. Изменение интервала времени между подачей заказа и осуществлением поставок представлено в таблице 3.2.

Таблица 3.2. Изменение интервала времени между подачей заказа и осуществлением поставок

Время поставки заказа, неделя	1	2	3	4
Вероятность	0,3	0,3	0,15	0,25

Единичная стоимость хранения запасов равна 50 коп. в неделю и рассчитывается для общего размера запаса, оставшегося на конец недели. Стоимость заказа - 60 руб., а отсутствие аккумуляторов на складе оценивается в 30 руб. неделю.

Используя имитационную модель для периода в 24 недели, оценить среднюю недельную стоимость проведения изложенной выше политики. Все расчеты производятся в начале недели, а подача заказов и поставки по ним - в конце недели.

Решим задачу имитационного моделирования управления запасами методом Монте-Карло.

Решение. Построим функцию распределения величины объема продаж в неделю и интервалы случайных чисел для значений стохастической переменной. Соответствующие значения указаны в четвертом и пятом столбцах таблицы 3.3. Расчеты произведены с использованием табличного процессора MS Excel. Вид листа Excel с формулами представлен на рисунке 3.1.

Таблица 3.3. Параметры стохастической переменной объем спроса

Спрос в неделю	**Частота**	**Вероятность**	**Значение функции распределения**	**Интервал случайных чисел**
670	3	0,03	0,03	от 1 до 3
675	7	0,07	0,1	от 4 до 10
680	5	0,05	0,15	от 11 до 15
685	10	0,1	0,25	от 16 до 25
690	13	0,13	0,38	от 26 до 38
695	8	0,08	0,46	от 39 до 46
700	5	0,05	0,51	от 47 до 51
705	2	0,02	0,53	от 52 до 53
710	13	0,13	0,66	от 54 до 66
715	10	0,1	0,76	от 67 до 76
720	9	0,09	0,85	от 77 до 85
725	3	0,03	0,88	от 86 до 88
730	7	0,07	0,95	от 89 до 95
735	2	0,02	0,97	от 96 до 97
740	3	0,03	1	от 98 до 100
Итого	100	1		

Рисунок 3.1. Вид листа Excel с расчетными формулами для определения параметров спроса

A Спрос в неделю	B Частота	C Вероятность	D Значение функции распределения	E F G H Интервал случайных чисел			
2 670	3	=B2/$B$17	=C2	от	1	ДО	=D2*100
3 675	7	=B3/SBS17	=D2+C3	от	=H2+1	ДО	=D3*100
4 680	5	=B4 SBS17	=D3+C4	от	=H3+1	ДО	=D4*100
5 685	10	=B5/$B$17	=D4+C5	от	=H4+1	ДО	=D5*100
6 690	13	=B6 SBS17	=D5+C6	от	=H5+1	ДО	=D6*100
7 695	8	=B7/$B$17	=D6+C7	от	=H6+1	ДО	=D7*100
8 700	5	=B8 SBS17	=D7+C8	от	=H7+1	ДО	=D8*100
9 705	2	=B9 SBS17	=D8+C9	от	=H8+1	ДО	=D9*100
10 710	13	=B 10 SBS17	=D9+C10	от	=H9+1	ДО	=D10*100
11 715	10	=B11 SBS17	=D10+C11	от	=H10+1	ДО	=D11*100
12 720	9	=B 12/SBS17	=D11+C12	от	=H11+1	ДО	=D12*100
13 725	3	=B13 SBS17	=D12+C13	от	=H12+1	ДО	=D13*100
14 730	7	=B14 SBS17	=D13+C14	от	=H13+1	ДО	=D14*100
15 735	2	=B 15/SBS17	=D14+C15	от	=H14+1	ДО	=D15*100
16 740	3	=B16 SBS17	=D15+C16	от	=H15+1	ДО	=D16*100
17 Итого	=СУММ(B2B 16)	=СУММ(C2:C 16)					

Аналогично построим функцию распределения и интервалы случайных чисел для времени выполнения поставок (таблица 3.4., рисунок 3.2.).

Таблица 3.4. Параметры стохастической переменной время поставок

Время поставок, мес.	**Вероятность**	**Значение функции распределения**	**Интервал случайных чисел**
1	0,3	0,3	от 1 до 30
2	0,3	0,6	от 31 до 60
3	0,15	0,75	от 61 до 75
4	0,25	1	от 76 до 100
Итого	1		

Рисунок 3.2. Вид листа Excel с расчетными формулами для определения стохастической переменной время поставок

J	K	1 L	м	N	0	Р
Время поставок, мес.	**Вероятность**	**Значение функции распределения**	**Интервал случайных чисел**			
1	**0,3**	**=К2**	**от**	|1	**ДО**	**=L2*100**
2	**0,3**	**=L2+K3**	**от**	=Р2г1	**ДО**	**=L3*100**
3	**0,15**	**=L3+K4**	**от**	=Р3+1	**ДО**	**=L4*100**
4	**0,25**	**=L4+K5**	**от**	=Р4*1	**ДО**	**=L5*100**
Итого	**=СУММ(К2Х5)**					

Процесс имитации реализуется в процессе выполнения четырех шагов:

1. Каждая имитируемая неделя начинается с проверки, поступил ли сделанный заказ. Если заказ выполнен, то текущий запас увеличивается на величину заказа (в данном случае — на 2500 шт.).

2. Путем выбора случайного числа генерируется недельный спрос для соответствующего распределения вероятностей.

3. Рассчитывается итоговый запас, равный исходному запасу за вычетом величины продаж. Если запас недостаточен для удовлетворения недельного спроса, спрос удовлетворяется, насколько это возможно.

Фиксируется число нереализованных продаж.

4. Определяется, снизился ли запас до точки восстановления (в примере — 1300 шт.). Если да, причем не ожидается поступления заказа, сделанного ранее, то делается заказ.

Для генерации случайных чисел воспользуемся формулой =СЛУЧМЕЖДУ(1;100) и результаты зафиксируем, так как эти числа могут изменяться со временем. Для определения спроса в зависимости от случайного числа воспользуемся функцией ЕСЛИ().

Таблица с результаты имитации представлены в таблице 3.5 - в расчетном виде и на рисунке 3.3.- в формульном виде.

Таблица 3.5. Результаты имитационного моделирования

Неделя	Поступление	Начальный запас	Случайное число	Спрос	Объем продаж	Конечный запас	Потери продаж	Делать заказ?	Случайное число	Время поставок
1	0	1800	36	690	690	1110	0	Да	66	3
2	0	1110	94	730	730	380	0	Нет		
3	0	380	56	710	380	0	330	Нет		
4	0	0	43	695	0	0	695	Нет		
5	2500	2500	12	680	680	1820	0	Нет		
6	0	1820	61	710	710	1110	0	Да	21	1
7	0	1110	26	690	690	420	0	Нет		
8	2500	2920	65	710	710	2210	0	Нет		
9	0	2210	61	710	710	1500	0	Нет		
10	0	1500	2	670	670	830	0	Да	75	3
11	0	830	9	675	675	155	0	Нет		
12	0	155	57	710	155	0	555	Нет		
13	0	0	2	670	0	0	670	Нет		
14	2500	2500	83	720	720	1780	0	Нет		
15	0	1780	65	710	710	1070	0	Да	57	2
16	0	1070	77	720	720	350	0	Нет		
17	0	350	95	730	350	0	380	Нет		
18	2500	2500	79	720	720	1780	0	Нет		
19	0	1780	90	730	730	1050	0	Да	51	2
20	0	1050	51	700	700	350	0	Нет		
21	0	350	86	725	350	0	375	Нет		
22	2500	2500	100	740	740	1760	0	Нет		
23	0	1760	98	740	740	1020	0	Да	34	2
24	0	1020	74	715	715	305	0	Да		
Итого						19000	3005			

Результаты имитационного эксперимента:

- конечный суммарный запас - 19000 штук;
- средний конечный запас 19000/24 =791,67 штук;
- число упущенных продаж - 3005;
- среднее число упущенных продаж 3005/24 = 125,2 шт. в месяц;
- за все время придется сделать 6 заказов;
- среднее число заказов 6/24 = 0,25 заказа в неделю;
- за весь период количество недель с упущенными продажами (отсутствие аккумуляторов на складе) составило 6;
- среднее число упущенных недель 6/24 = 0,25 недель.

Определим среднюю стоимость проведения изложенной выше политики
в неделю. Для этого вычислим ее составляющие:

Еженедельная стоимость заказов = Затраты на один заказ х Среднее число заказов в неделю = 60 • 0,25 = 15 руб.

Еженедельная стоимость хранения = Затраты на хранение одной единицы в течение недели х Средняя величина конечного запаса = 0,5^ 91,67 = 395,83руб.

Еженедельная стоимость упущенных продаж = Стоимость упущенной продажи х Среднее число упущенных продаж в неделю = 30 • 0,25 = 7,5 руб.

Таким образом,

Общая еженедельная стоимость = Стоимость заказов + Стоимость хранения + Стоимость упущенных продаж = 15 + 395,83+ 7,5= 418,33 руб.

Вывод. Проведенный эксперимент показывает, что за 24 недели придется сделать 6 заказов, общая еженедельная стоимость составит 418,33 руб.

C	D	E	F	G	H	I	J	K	L	I M
1										
Неделя	Поступлен	Начальный запас	Случайное	Спрос	Объем продаж	Конечный запас	Потери продаж	Делать заказ?	Случайн	Время поставок
3 1	0	1800	36	=ЕСЛИ(И(1	=ЕСМ(Е3>=G3;G3;Е	=Е3-Н3	=ЕСТН(Н3<Е3;G3-	Да	66	=ЕСЖ(ИаЗ>=Лист2
4 р	0	=13	94	=ЕСЛИ(И(1	=ЕСМ(Е4>=G4;G4;Е	=Е4-Н4	=ЕСМ(Н4<Е4;G4-	Нет		
5 3	0	=14	56	=ЕСЛИ(И(1	=ЕС ЛИ(Е5 >=G5	=Е5-Н5	=ЕСJm(Н5<G5;G5-	Нет		
б 4	0	=D6+I5	43	=ЕСЛИ(И(1	=ЕСJm(Е6>=G6;G6;	=Е6-Н6	=ЕСМ(Н6<G6;G6-	Нет		
7 5	2500	=D7+I6	12	=ЕСЛИ(И(1	=ЕС JIН(Е7>=G7	=Е7-Н7	=ЕСЖ(Н7 <G7 ;G7 -	Нет		
8 6	0	=D8+I7	61	=ЕСЛИ(И(1	=ЕСМ(Е8>=G8;G8;Е	=Е8-Н8	=ЕСЖ(Н8<С8;С8-	Да	21	=ЕСJ1Н(Н(L8>=JIнст2
9 7	0	=D9+I8	26	=ЕСЛИ(И(1	=ЕС„ТН(Е9>=G9;G9	=Е9-Н9	=ЕСЖ(Н9<С9;О9-	Нет		
10 8	2500	=D	65	=ЕСЛИ(И(1	=ЕС ЛИ(Е 10>=G	=Е10-	=ЕСЖ(Н 10<G	Нет		
И 9	0	=D11+1	61	=ЕСЛИ(И((	=ЕСJМ(Е11>=G11;G	=Е11-	=ЕСЖ(Н11<G11;G11-	Нет		
12 10	0	=D12+I	2	=ЕСЛИ(И(1	=ЕСЖ(Е 12>=G	=Е12-	=ЕСЖ(Н12<С12;С12-	Да	75	=ЕСЖ(И(1.12>=Лист2! SN$2;L
13 И	0	=4)13+1	9	=ЕСЛИ(И(1	=ЕС ЛИ(Е 13>=G	=Е13-	=ЕСJIН(Н13<G13;G1	Нет		
14 12	0	=D14+I	57	=ЕСЛИ(И(1	=ЕС ЛЩЕ 14>=G	=Е14-	=ЕСJIН(Н14<G14;G1	Нет		
15 13	0	=D	2	=ЕСЛИ(И(1	=ЕС ЛИ(Е 15>=G 15	=Е15-	=ЕС^Н(Н15<G15;G15	Нет		
16 14	2500	=D16+I	83	=ЕСЛИ(И(1	=ЕСЖ(Е 16>=G	=Е16-	=ЕСЖ(Н 16<G 16;G	Нет		
17 15	0	=D17+I	65	=ЕСЛИ(И(1	=ЕС ЛЩЕ 17>=G	=Е17-	=ЕСЖ(Н17<Ш7;С17-	Да	57	=ЕС ЛИ(И(117>=Лист2! SNS2 ;L
18 16	0	=D18+I	77	=ЕСЛИ(И(1	=ЕС ЛИ(Е 18>=G	=Е18-	=ЕСЖ(Н18<С18;Ш8-	Нет		
19 17	0	=D19+I	95	=ЕСЛИ(И(1	=ЕС ЛИ(Е 19>=G	=Е19-	=ЕСЖ(Н 19<G	Нет		
20 18	2500	=D20+I	79	=ЕСЛИ(И(1	=ЕСJМ(Е20>=G20;G	=Е20-	=ЕСЖ(Н20<020:020-	Нет		
21 19	0	=D21+1	90	=ЕСЛИ(И(1	=ЕСЛИ(Е21>=G21	=Е21-	=ЕСЖ(Н21 <G21	Да	51	=ЕСЖ(И(Ь21 >=Лист2! SNS2 ;L21
22 20	0	=D22+I	51	=ЕСЛИ(И(1	=ЕСМ(Е22>=G22;G	=Е22-	=ЕСЖ(Н22<О22;О22-	Нет		
23 21	0	=D23+I	86	=ЕСЛИ(И(1	=ЕС НН(Е23>=G23	=Е23-	=ЕСЖ(Н23<О23;С23-	Нет		
*ЖЛ*___	2500,	=D24+I	100	=ЕСЛИ(И(1	=ЕС,[1]М(Е24>=G24:	=Е24-	=ЕСЖ(Н24<О24:О24-	____Нет____		

Рисунок 3.3. Таблица результатов имитационного моделирования в формульном виде

Контрольные вопросы

1. Что такое имитационное моделирование?
2. В чем заключается метод Монте-Карло?
3. С помощью чего при использовании метода Монте-Карло генерируются результаты наблюдений?

ПРАКТИЧЕСКОЕ ЗАНЯТИЕ №3
ПОСТРОЕНИЕ ПРОСТРАНСТВЕННОЙ МОДЕЛИ РЕГРЕССИИ

Пример 1. Построение пространственной модели связи объёма реализации одного из продуктов фирмы от нескольких факторных признаков. Пусть имеются некоторые данные об объеме реализации одного из продуктов фирмы. На основании содержательного анализа составлен перечень показателей, которые предполагается включить в модель, и составлена таблица исходных данных (табл. 1.). Задача решается с помощью методов корреляционного и регрессионного анализа.

1. Необходимо составить матрицу парных коэффициентов корреляции и на ее основе дать рекомендации о включении в модель тех или иных факторов.

Решение. Парные коэффициенты корреляции вычисляются на основе формул (1) и (2):

$$r_{xy} == \frac{\overline{yx} - \overline{y} \cdot \overline{x}}{\sigma_x \sigma_y}, \qquad (1),$$

$$\sigma_x = \sqrt{\overline{x^2} - \overline{x}^2}. \qquad (2).$$

Для того, чтобы вычислить, например, коэффициент корреляции между x_2 и x_3, формулы необходимо записать в следующем виде:

$$r_{x_2 x_3} == \frac{\overline{x_2 x_3} - \overline{x_2} \cdot \overline{x_3}}{\sigma_{x_2} \sigma_{x_3}},$$

$$\sigma_{x_2} = \sqrt{\overline{x_2^2} - \overline{x_2}^2},$$

$$\sigma_{x_3} = \sqrt{\overline{x_3^2} - \overline{x_3}^2}.$$

Таблица 1.

Объем реализации	Время	Реклама	Цена	Цена конкурента	Индекс потребит. расходов
y	x_1	x_2	x_3	x_4	x_5
126	1	4	15,0	17,0	100,0
137	2	4,8	14,8	17,3	98,4
148	3	3,8	15,2	16,8	101,2
191	4	8,7	15,5	16,2	103,5
274	5	8,2	15,5	16,0	104,1

370	6	9,7	16,0	18,0	107,0
432	7	14,7	18,1	2,02	107,4
445	8	18,7	13,0	15,8	108,5
367	9	19,8	15,8	18,2	108,3
367	10	10,6	16,9	16,8	109,2
321	11	8,6	16,3	17,0	110,1
307	12	6,5	16,1	18,3	110,7
331	13	12,6	15,4	16,4	110,3
345	14	6,5	15,7	16,2	111,8
364	15	5,8	16,0	17,7	112,3
384	16	5,7	15,1	16,2	112,9

Копируем таблицу в EXCEL.

Arial 14 Ж К Ч

A1 fx Объем

	A	B	C	D	E	F
1–3	Объем реализации	Время	Реклама	Цена	Цена кон-курента	Индекс потребит. расходов
4	y	x_1	x_2	x_3	x_4	x_5
5	126	1	4	15	17	100
6	137	2	4,8	14,8	17,3	98,4
7	148	3	3,8	15,2	16,8	101,2
8	191	4	8,7	15,5	16,2	103,5
9	274	5	8,2	15,5	16	104,1
10	370	6	9,7	16	18	107
11	432	7	14,7	18,1	2,02	107,4
12	445	8	18,7	13	15,8	108,5
13	367	9	19,8	15,8	18,2	108,3
14	367	10	10,6	16,9	16,8	109,2
15	321	11	8,6	16,3	17	110,1
16	307	12	6,5	16,1	18,3	110,7
17	331	13	12,6	15,4	16,4	110,3
18	345	14	6,5	15,7	16,2	111,8
19	364	15	5,8	16	17,7	112,3
20	384	16	5,7	15,1	16,2	112,9
21						

Рис.1. Исходные данные задачи

воспользуемся надстройкой «Анализ данных» пакета анализа EXCEL.

(см. «EXCEL в науке и практике» с.122.). Команда меню «Данные/ Анализ данных/Корреляция» (рис.2).

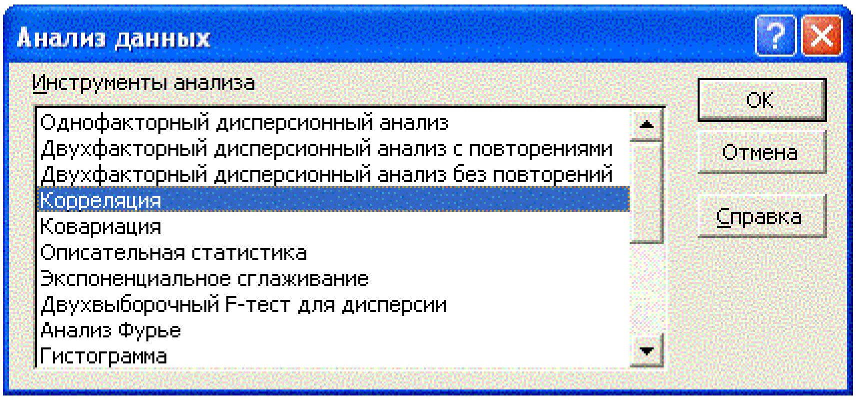


Рис. 2. Меню надстройки «Анализ данных»

В появившемся окне «Корреляция» ввести диапазон данных «у» и все «х» (рис.2.).

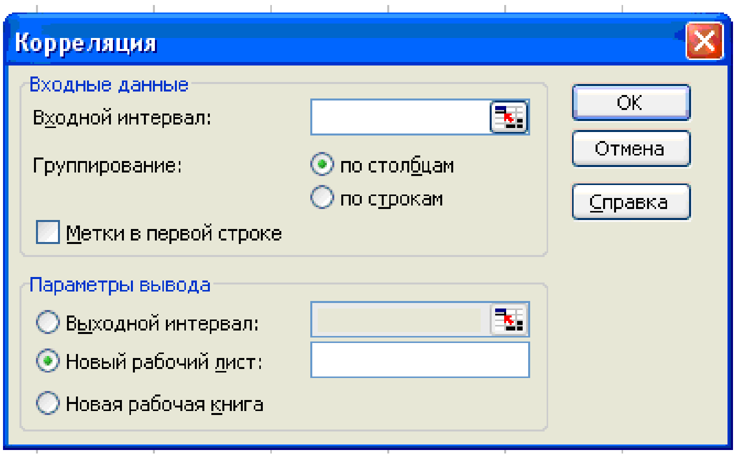


Рис. 3. Окно «Корреляция»

Для этого нужно в окошке входной интервал ввести диапазон данных A4:F20 и так как в сроке A4:F4 текстовые переменные установить птичку в окошке метки в первой строке.

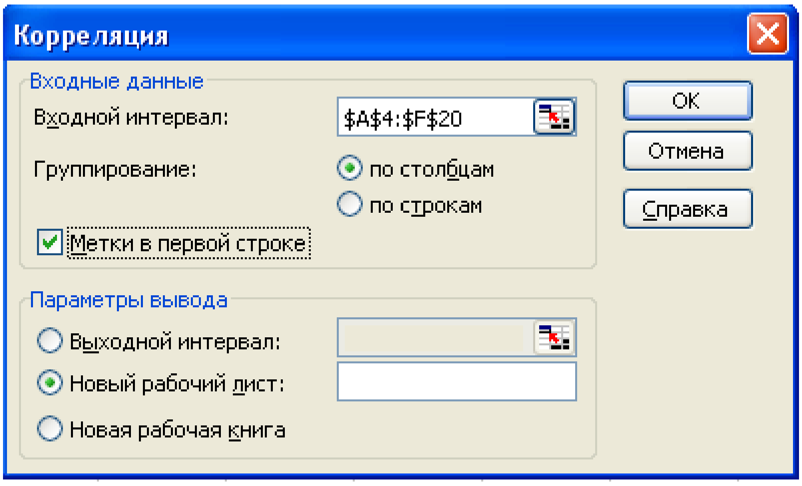


Рис. 3. Окно корреляция с введенным диапазоном данных

Нажать ОК. На новом листе появится матрица парных коэффициентов корреляции (корреляционная матрица). Первый столбец которой указывает на степень корреляционной связи факторов x1-x5 с результативным признаком y. Остальные столбцы указывают на корреляционную зависимость факторов между собой (рис.4.).

	y	*x1*	*x2*	*x3*	*x4*	*x5*
y	1					
x1	0,67798	1				
x2	0,64592	0,10646	1			
x3	0,2329	0,17372	-0,0034	1		
x4	-0,3195	0,08497	-0,289	-0,5183	1	
x5	0,81602	0,9602	0,27337	0,23543	-0,0027	1

Рис.4. Матрица парных коэффициентов корреляции

Как видно из рис.4., парный коэффициент корреляции факторов x1-x5 равен 0,96 ($r_{x_1x_5} = 0{,}96$), что свидетельствует о наличии мультиколлинеарности. Одним из условий регрессионной модели является предположение о линейной независимости объясняющих переменных (факторных признаков), т.е. решение задачи возможно только тогда, когда столбцы и строки исходных данных линейно независимы. Для экономических показателей это условие выполняется не всегда. Линейная или близкая к ней связь между факторами называется мультиколлинеарностью и приводит к линейной зависимости нормальных уравнений, что делает вычисление параметров либо невозможным, либо затрудняет содержательную интерпретацию параметров

Из этих двух переменных оставим в модели x_5 – индекс потребительских доходов, т.к. $r_{yx_5} > r_{yx_1}$. Да и экономический смысл подсказывает, что в индексе потребительских расходов информации больше, так как он также является функцией времени.

Для этого скопируем табл.1. на новый лист. И на нём удалим столбец с переменной x1. Чтобы удалить столбец с переменной x1 нужно щелкнуть правой кнопкой мыши на названии столбца А и в появившемся меню выбрать – удалить (рис.5.).

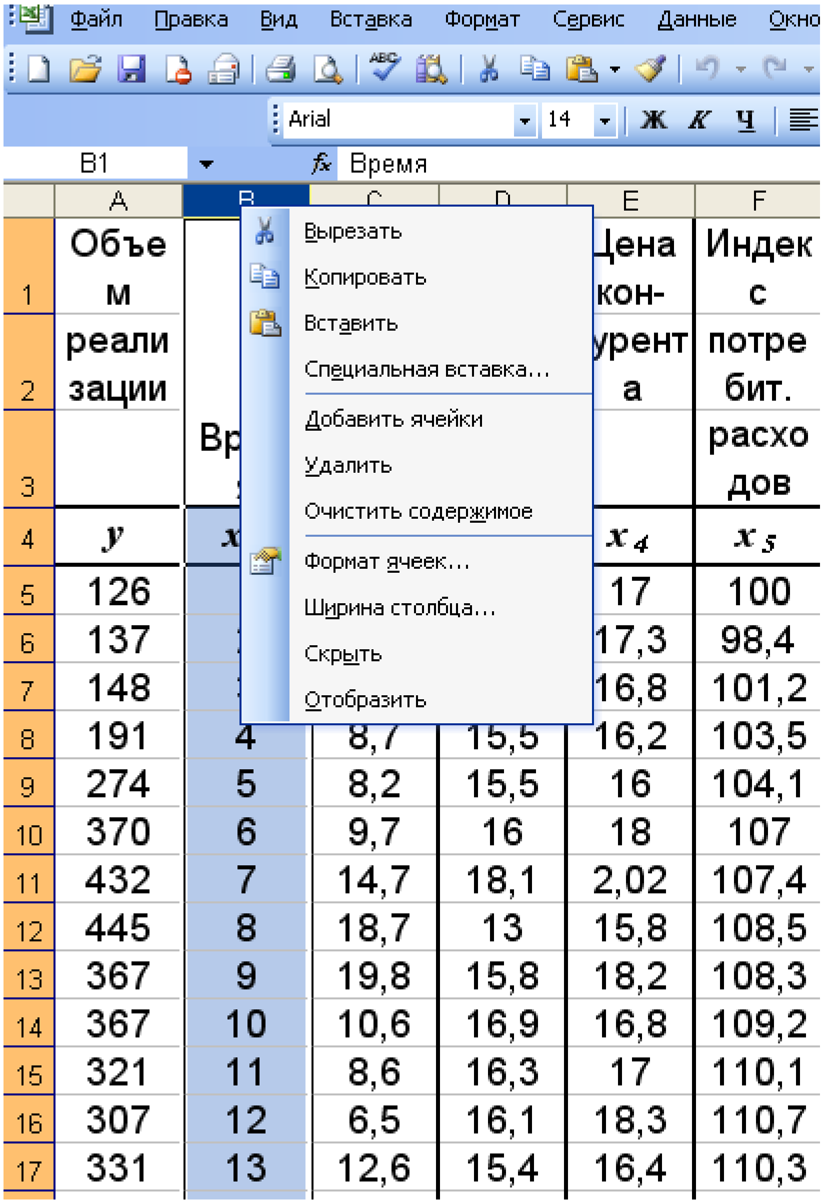


Рис.5. Удаление столбца с фактором x1

Получим табл.2.

Теперь можно преступить к регрессионному анализу. Команда «Данные/Анализ данных/Регрессия». Появится диалоговое окно «Регрессия» в котором нужно ввести входные интервалы x и y и установить птички в окошках «Метки» и «Уровень надёжности 0,95». Установка уровня надёжности 0,95 свидетельствует о выборе уровня значимости 5%.

Таблица 2.

Объем реализации	Реклама	Цена	Цена конкурента	Индекс потребительских расходов
y	$\boldsymbol{x_2}$	$\boldsymbol{x_3}$	$\boldsymbol{x_4}$	$\boldsymbol{x_5}$
126	4	15	17	100
137	4,8	14,8	17,3	98,4
148	3,8	15,2	16,8	101,2
191	8,7	15,5	16,2	103,5
274	8,2	15,5	16	104,1
370	9,7	16	18	107
432	14,7	18,1	2,02	107,4
445	18,7	13	15,8	108,5
367	19,8	15,8	18,2	108,3
367	10,6	16,9	16,8	109,2
321	8,6	16,3	17	110,1
307	6,5	16,1	18,3	110,7
331	12,6	15,4	16,4	110,3
345	6,5	15,7	16,2	111,8
364	5,8	16	17,7	112,3
384	5,7	15,1	16,2	112,9

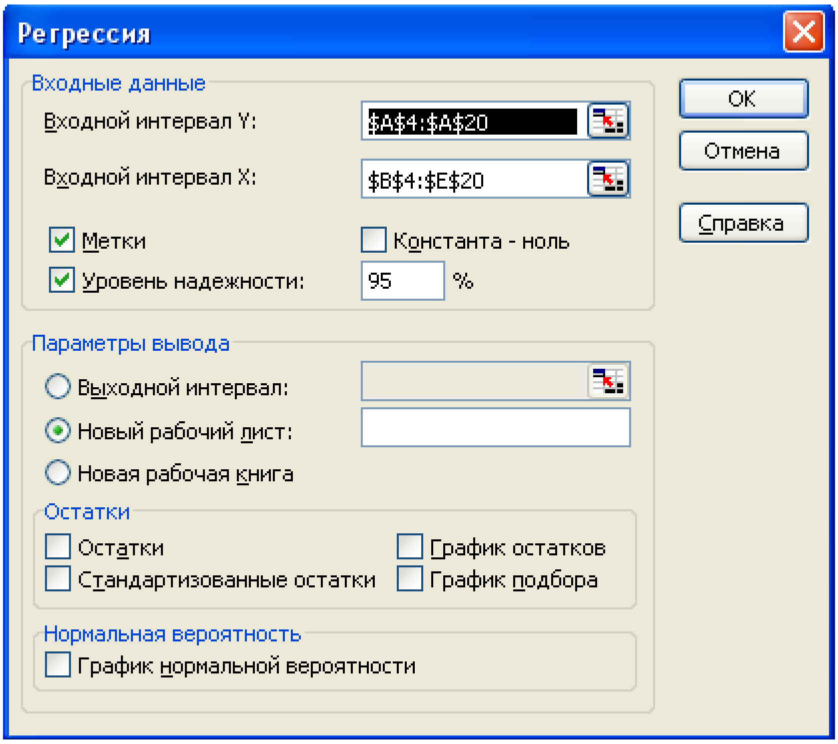


Рис.6. Окно «Регрессия» программы «Анализ данных»

После нажатия ОК, если входные интервалы заданы корректно появится Рис. 7. «Вывод итогов регрессии».

	A	B	C	D	E	F
1	ВЫВОД ИТОГОВ					
2						
3	*Регрессионная статистика*					
4	Множественный	0,948297672				
5	R-квадрат	0,899268474				
6	Нормированный	0,862638828				
7	Стандартная оши	38,12413985				
8	Наблюдения	16				
9						
10	Дисперсионный анализ					
11		*df*	*SS*	*MS*	*F*	*Значимость F*
12	Регрессия	4	142730,4871	35682,62177	24,55029123	1,95719E-05
13	Остаток	11	15987,95044	1453,45004		
14	Итого	15	158718,4375			
15						
16		*Коэффициенты*	*Стандартная ошибка*	*t-статистика*	*P-Значение*	*Нижние 95%*
17	Y-пересечение	-1346,771239	273,0636236	-4,932078542	0,000448185	-1947,780222
18	x2	7,907794525	2,27951453	3,469069586	0,005248568	2,890616876
19	x3	-5,887164166	11,65436616	-0,505146662	0,623423124	-31,53825112
20	x4	-6,483329464	3,331909533	-1,945829981	0,077667128	-13,8168129
21	x5	16,56164269	2,42642017	6,825546082	2,85523E-05	11,2211279
22						

Рис. 7. Вывод итогов регрессии

Как видно из рис.7., полученные итоги характеризуют линейную регрессию вида:

$$y = a_0 + a_2 x2 + a_3 x3 + a_4 x4 + a_5 x5$$

Только, прежде чем воспользоваться полученными результатами, необходимо проверить на значимость значения коэффициентов $a_0 - a_5$, а также проверить на значимость само уравнение регрессии. Проверить на значимость коэффициенты уравнения регрессии означает, что нужно убедиться в том, что значения коэффициентов получены неслучайно и им можно доверять. Другими словами, если коэффициенты уравнения регрессии значимы, то на их основе можно делать прогноз будущих периодов. Осуществляется такая проверка с помощью t-теста Стьюдента. Рассчитывается значение t-статистики и сравнивается с порогом. Как видно из итогов регрессии (рис.7), мы имеем значения соответствующих коэффициентов уравнения регрессии (столбец-коэффициенты), значения t-статистики для каждого коэффициента (столбец t-статистика), значения стандартных ошибок. Нет только пороговых значений для критерия Стьюдента. Их можно найти по таблицам для критерия Стьюдента или же воспользоваться p-значением, которое не должно превышать 0,05 при 5% уровне значимости. Тогда и соответствующий коэффициент будет значим.

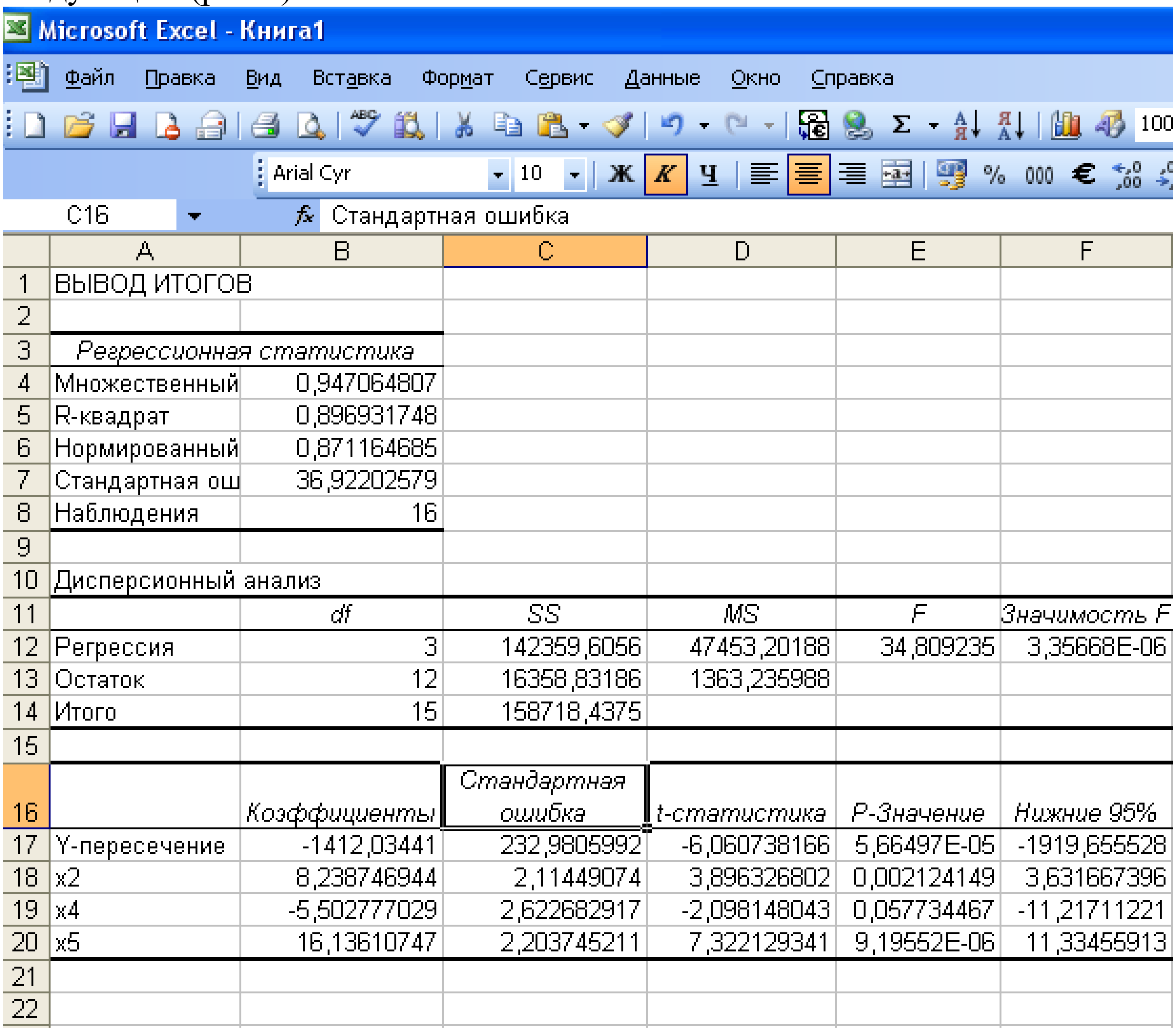
После этого нужно провести проверку на значимость выбранного уравнения регрессии. Она осуществляется с помощью F – теста Фишера. Рассчитывается F статистика и сравнивается с порогом. Пороговое значение выбирается по таблицам F распределения Фишера. В программе *F*, а можно воспользоваться полем «Значимость F». При 5% уровне значимости его значение должно быть меньше 0,05 для того, чтобы выбранное уравнение регрессии было значимым. У нас «Значимость F» $=1{,}96*10^{-5}$ т.е. выбранное нами линейное уравнение регрессии значимо.

Таблица3.

Объем реализации	Реклама	Цена конкурента	Индекс потребительских расходов
y	x_2	x_4	x_5
126	4	17	100
137	4,8	17,3	98,4
148	3,8	16,8	101,2
191	8,7	16,2	103,5
274	8,2	16	104,1
370	9,7	18	107
432	14,7	2,02	107,4
445	18,7	15,8	108,5
367	19,8	18,2	108,3
367	10,6	16,8	109,2
321	8,6	17	110,1
307	6,5	18,3	110,7
331	12,6	16,4	110,3
345	6,5	16,2	111,8
364	5,8	17,7	112,3
384	5,7	16,2	112,9

Из рис.7. видно, что коэффициент a_3 при переменной x3 незначим, так как P-значение=0,62. Следовательно переменную x3 нужно также исключить из модели. Чтобы сохранить промежуточные таблицы, опять копируем табл.2. на лист 3. и удаляем столбец с x3. Получим таблицу 3.

Опять вычисляем регрессию с помощью программы «Анализ данных». Окно программы «Регрессия» будет таким же как на рис.6. только изменится блок ячеек «Входной интервал x». Вывод итогов будет следующим (рис.8).



	A	B	C	D	E	F
1	ВЫВОД ИТОГОВ					
2						
3	*Регрессионная статистика*					
4	Множественный	0,947064807				
5	R-квадрат	0,896931748				
6	Нормированный	0,871164685				
7	Стандартная ош	36,92202579				
8	Наблюдения	16				
9						
10	Дисперсионный анализ					
11		*df*	*SS*	*MS*	*F*	*Значимость F*
12	Регрессия	3	142359,6056	47453,20188	34,809235	3,35668E-06
13	Остаток	12	16358,83186	1363,235988		
14	Итого	15	158718,4375			
15						
16		*Коэффициенты*	*Стандартная ошибка*	*t-статистика*	*P-Значение*	*Нижние 95%*
17	Y-пересечение	-1412,03441	232,9805992	-6,060738166	5,66497E-05	-1919,655528
18	x2	8,238746944	2,11449074	3,896326802	0,002124149	3,631667396
19	x4	-5,502777029	2,622682917	-2,098148043	0,057734467	-11,21711221
20	x5	16,13610747	2,203745211	7,322129341	9,19552E-06	11,33455913
21						
22						

Рис.8. Вывод итогов регрессии

Поскольку a_4 всё еще остаётся незначимым, то исключаем x4. Окончательные итоги регрессии приведены на рис.9. Для повышения достоверности полученных результатов, оба теста повторяются с уровнем значимости в 1%. Он устанавливается с помощью уровня надёжности 0,99. В нашем случае коэффициенты уравнения регрессии a_2 и a_5 , а также само уравнение регрессии будут значимы.

	A	B	C	D	E	F
1	ВЫВОД ИТОГОВ					
2						
3	*Регрессионная статистика*					
4	Множественны	0,926887773				
5	R-квадрат	0,859120944				
6	Нормированны	0,837447243				
7	Стандартная о	41,47297887				
8	Наблюдения	16				
9						
10	Дисперсионный анализ					
11		*df*	*SS*	*MS*	*F*	*Значимость F*
12	Регрессия	2	136358,3338	68179,1669	39,6388667	2,93428E-06
13	Остаток	13	22360,10369	1720,007977		
14	Итого	15	158718,4375			
15						
16		*Коэффициенты*	*Стандартная ошибка*	*t-статистика*	*P-Значение*	*Нижние 95%*
17	Y-пересечение	-1471,314319	259,7660084	-5,663998643	7,7458E-05	-2032,504661
18	x2	9,568413823	2,265936415	4,222719472	0,00099649	4,673155824
19	x5	15,75287403	2,466858435	6,385803824	2,3963E-05	10,42355039

Рис. 9. Вывод итогов регрессии

В результате получим следующую пространственную регрессионную модель.

$$y = -1471{,}31 + 9{,}57x2 + 15{,}75x5$$

Экономический смысл полученного уравнения регрессии. Если в рекламу вложить миллион рублей, то выручка от реализации продукта увеличится на 9.57 миллионов рублей. Вот почему мы каждый фильм по телевизору смотрим по 3-4 часа. Увеличение индекса потребительских расходов населения на 10% увеличивает выручку от реализации продукта на 1,6 миллиона руб.

По аналогичной методике постройте модель оценки производительности труда по относительным показателям.

Контрольные вопросы

1. Сущность регрессионного анализа.
2. Что такое пространственная модель регрессии?
3. Области применения регрессионного анализа.
4. Реализация пространственной модели регрессии с помощью программных продуктов.
5. Какая формула используется при расчете t-статистики через коэффициент детерминации для оценки уравнения множественной регрессии?

ПРАКТИЧЕСКОЕ ЗАНЯТИЕ № 4
МОДЕЛИРОВАНИЕ ОДНОМЕРНЫХ ВРЕМЕННЫХ РЯДОВ. ПОСТРОЕНИЕ АДДИТИВНОЙ МОДЕЛИ ПОТРЕБЛЕНИЯ ЭЛЕКТРОЭНЕРГИИ

Краткие сведения из теории

Моделирование сезонных и циклических колебаний

Существует несколько подходов к анализу структуры динамических рядов, содержащих сезонные или циклические колебания. Поскольку моделирование циклических колебаний в целом осуществляется аналогично моделированию сезонных колебаний, мы рассмотрим только методы моделирования последних.

Простейший подход – расчет значений сезонной компоненты ***методом скользящей средней*** и построение аддитивной или мультипликативной модели временного ряда.

Общий вид ***аддитивной модели*** следующий:

$Y = T + S + E$.

Эта модель предполагает, что каждый уровень временного ряда может быть представлен как сумма трендовой (*T*), сезонной (*S*) и случайной (*E*) компонент.

Общий вид ***мультипликативной модели*** выглядит так:

$Y = T \cdot S \cdot E$.

Эта модель предполагает, что каждый уровень временного ряда может быть представлен как произведение трендовой (*T*), сезонной (*S*) и случайной (*E*) компонент.

Выбор одной из двух моделей осуществляется на основе анализа структуры сезонных колебаний. Если амплитуда колебаний приблизительно постоянна, строят аддитивную модель временного ряда, в которой значения сезонной компоненты предполагаются постоянными для различных циклов. Если амплитуда сезонных колебаний возрастает или уменьшается, строят мультипликативную модель временного ряда, которая ставит уровни ряда в зависимость от значений сезонной компоненты.

Построение аддитивной и мультипликативной моделей сводится к расчету значений *T, S* и *E* для каждого уровня ряда.

Процесс построения модели включает в себя следующие шаги:

Шаг 1. Выравнивание исходного ряда методом скользящей средней.

Шаг 2. Расчет значений сезонной компоненты S.

Шаг 3. Устранение сезонной компоненты из исходных уровней ряда и получение выравненных данных (*T* + *E*) в аддитивной или ($T \cdot E$) в мультипликативной модели.

Шаг 4. Аналитическое выравнивание уровней ($T + E$) или ($T \cdot E$) и расчет значений T с использованием полученного уравнения тренда.

Шаг 5. Расчет полученных по модели значений ($T + E$) или ($T \cdot E$).

Шаг 6. Расчет абсолютных и/или относительных ошибок.

Если полученные значения ошибок не содержат автокорреляции, ими можно заменить исходные уровни ряда и в дальнейшем использовать временной ряд ошибок E для анализа взаимосвязи исходного ряда и других временных рядов.

Подробнее методику построения каждой из моделей рассмотрим на примерах.

Пример 1. Построение аддитивной модели временного ряда.

Обратимся к данным об объеме потребления электроэнергии жителями района за последние четыре года (табл. 4.10).

Таблица 4.10.

Год	№ квартала	Потребление эл/энергии	Итого за чктыре квартала	Скользящая средняя за четыре квартала	Центрированная скользящая средняя	Оценка сезонной компоненты
	1	2	3	4	5	6
1	1	6,0	---	---	---	---
	2	4,4	24,4	6,10	---	---
	3	5,0	25,6	6,40	6,250	-1,250
	4	9,0	26,0	6,50	6,450	2,550
2	5	7,2	27,0	6,75	6,625	0,575
	6	4,8	28,0	7,00	6,875	-2,075
	7	6,0	28,8	7,20	7,100	-1,100
	8	10,0	29,6	7,40	7,300	2,700
3	9	8,0	30,0	7,50	7,450	0,550
	10	5,6	31,0	7,75	7,625	-2,025
	11	6,4	32,0	8,00	7,875	-1,475
	12	11,0	33,0	8,25	8,125	2,875
4	13	9,0	33,6	8,40	8,325	0,675
	14	6,6	33,4	8,35	8,375	-1,775
	15	7,0	---	---	---	---
	16	10,8			---	---

Построение аддитивной моделисводится к расчету значений T, S и E для каждого уровня ряда.

Процесс построения модели включает в себя следующие шаги:

Шаг 1. Проведем выравнивание исходных уровней ряда методом скользящей средней. Для этого:

- просуммируем уровни ряда последовательно за каждые четыре квартала со сдвигом на один момент времени, и определим условные годовые объемы потребления электроэнергии (гр. 3 табл. 4.10):

6,0 + 4,4 + 5,0 + 9,0 = 24,4 ,

4,4 + 5,0 + 9,0 +7,2 = 25,6 и т.д.;

- разделив полученные суммы на четыре, найдем скользящие средние (гр. 4 табл. 4.10). Отметим, что полученные таким образом выравненные значения уже не содержат сезонной компоненты;
- приведем эти значения в соответствие с фактическими моментами времени, для чего найдем средние из двух последовательных скользящих средних – центрированные скользящие средние (гр. 5 табл. 4.10):

$$(6{,}10 + 6{,}40) : 2 = 6{,}250\ ,$$

$$(6{,}40 + 6{,}50) : 2 = 6{,}450 \text{ и т.д.}$$

Шаг 2. Найдем оценки сезонной компоненты как разность между фактическими уровнями ряда и центрированными скользящими средними (гр. 6 = гр. 2 – гр. 5):

$$5{,}0 – 6{,}250 = -1{,}250\ ,$$

$$9{,}0 – 6{,}450 = 2{,}550 \text{ и т.д.}$$

Используем эти оценки для расчета значений сезонной компоненты S (табл. 4.11). Для этого найдем средние за каждый квартал (по всем годам) оценки сезонной компоненты S_i. В моделях с сезонной компонентой обычно предполагается, что сезонные воздействия за период взаимопогашаются. *В аддитивной модели взаимопогашаемость выражается в том, что сумма значений сезонной компоненты по всем кварталам должна быть равна нулю.*

Показатели	Год	№ квартала, i			
		I	II	III	IV
	1	---	---	-1,250	2,550
	2	0,575	-2,075	-1,100	2,700
	3	0,550	-2,025	-1,475	2,875
	4	0,675	-1,775	---	---
Итого за i-ый квартал (за все годы)		1,800	-5,875	-3,825	8,125
Средняя оценка сезонной компоненты для i-го квартала, $\bar{S}_i$		0,600	-1,958	-1,275	2,708
Скорректированная сезонная компонента, S_i		0,581	-1,977	-1,294	2,690

Для данной модели имеем:

$$0{,}6 – 1{,}958 – 1{,}275 + 2{,}708 = 0{,}075\ .$$

Определим корректирующий коэффициент:

$$k = \frac{0{,}075}{4} = 0{,}01875\ .$$

Рассчитаем скорректированные значения сезонной компоненты как разность между ее средней оценкой и корректирующим коэффициентом k:

$$S_i = \overline{S}_i - k, \quad i = \overline{1,4}.$$

Проверим условие равенства нулю суммы значений сезонной компоненты:

$$0,581 - 1,977 - 1,294 + 2,690 = 0 .$$

Таким образом, получены следующие значения сезонной компоненты:

I квартал: $S_1 = 0,581$;

II квартал: $S_2 = -1,979$;

III квартал: $S_3 = -1,294$;

IV квартал: $S_4 = 2,690$.

Занесем полученные значения скорректированной сезонной компоненты в табл. 4.11 для соответствующих кварталов каждого года (стр. 3).

Шаг 3. Устраним (элиминируем) влияние сезонной компоненты, вычитая ее значения из каждого уровня исходного временного ряда. Получим величины $T + E = Y - S$ (гр.4. табл. 4.12). Эти значения рассчитываются за каждый момент времени и содержат только тенденцию и случайную компоненту.

Шаг 4. Определим компоненту *T* данной модели. Для этого проведем аналитическое выравнивание ряда (*T* + *E*) с помощью линейного тренда. Результаты аналитического выравнивания следующие:

Константа (*a*)	5,715416
Коэффициент регрессии (*b*)	0,186421
Стандартная ошибка коэффициента регрессии	0,015488
R-квадрат	0,914971
Число наблюдений	16
Число степеней свободы	14

Таким образом, имеем следующий линейный тренд:

$$T = 5,7155 + 0.1864 \cdot t \, .$$

Подставляя в это уравнение значения $t = 1, \ldots, 16$, найдем уровни Т для каждого момента времени (гр. 5 табл. 4.12).

t	y_t	S_i	$T+E=$ $=y_t-S_t$	T	$T+S$	$E=y_t-$ $-(T+S)$	E^2
1	2	3	4	5	6	7	8
1	6,0	0,581	5,419	5,9019	6,4829	-0,4829	0,2332
2	4,4	-1,977	6,377	6,0883	4,1113	0,2887	0,0833
3	5,0	-1,294	6,294	6,2747	4,9807	0,0193	0,0004
4	9,0	2,690	6,310	6,4611	9,1511	-0,1511	0,0228
5	7,2	0,581	6,619	6,6476	7,2286	-0,0286	0,0008
6	4,8	-1,977	6,777	6,8340	4,8570	-0,0570	0,0032
7	6,0	-1,294	7,294	7,0204	5,7264	0,2736	0,0749
8	10,0	2,690	7,310	7,2068	9,8968	0,1032	0,0107
9	8,0	0,581	7,419	7,3932	7,9742	0,0258	0,0007
10	5,6	-1,977	7,577	7,5796	5,6026	-0,0026	0,0000
11	6,4	-1,294	7,694	7,7660	6,4720	-0,0720	0,0052
12	11,0	2,690	8,310	7,9524	10,6424	0,3576	0,1278
13	9,0	0,581	8,419	8,1389	8,7199	0,2801	0,0785
14	6,6	-1,977	8,577	8,3253	6,3483	0,2517	0,0634
15	7,0	-1,294	8,294	8,5117	7,2177	-0,2177	0,0474
16	10,8	2,690	8,110	8,6981	11,3881	-0,5881	0,3458

График уравнения тренда приведен на рис. 4.4.

Шаг 5. Найдем значения уровней ряда, полученные по аддитивной модели. Для этого прибавим к уровням T значения сезонной компоненты для соответствующих кварталов. Графически значения (T + S) представлены на рис. 4.4.

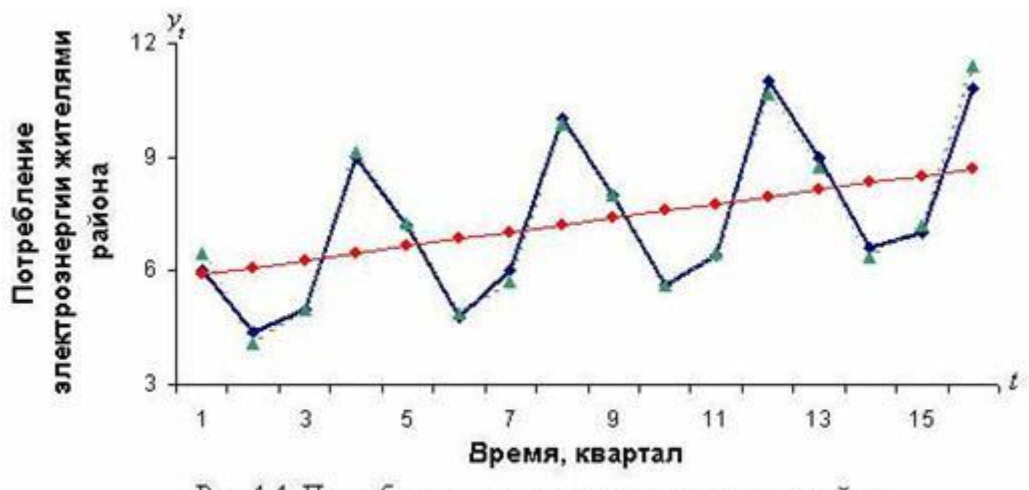


Рис.4.4. Потребление элктроэнергии жителями района (фактические, выравненные и полученные по аддитивной модели значения уровней ряда):
фактические значения;
тренд T;
значения $(T+S)$.

Шаг 6. В соответствии с методикой построения аддитивной модели расчет ошибки производится по формуле

$$E = Y - (T + S).$$

Это абсолютная ошибка. Численные значения абсолютных ошибок приведены в гр. 7 таблицы 4.12.

По аналогии с моделью регрессии для оценки качества построения модели или для выбора наилучшей модели можно применять сумму квадратов полученных абсо-лютных ошибок. Для данной аддитивной модели сумма квадратов абсолютных ошибок равна 1,0981. По отношению к общей сумме квадратов отклонений уровней ряда от своего среднего уровня, равной 67,120, эта величина составляет чуть более 1,6%:

$$100 - \frac{1 - 1{,}0981}{67{,}120} \cdot 100 = 1{,}6361.$$

Следовательно, можно сказать, что аддитивная модель объясняет 98,3639% общей вариации уровней временного ряда потребления электроэнергии за последние 16 кварталов.

Пример 2. *Прогнозирование по аддитивной модели*.

Предположим, по данным примера 4.4 требуется дать прогноз потребления электроэнергии жителями района в течение первого полугодия ближайшего следующего года.

Прогнозное значение F_i уровня динамического ряда в аддитивной модели в соответствии с соотношением $Y = T + S + E$ есть сумма трендовой и сезонной компонент.

Объем электроэнергии, потребленной в течение первого полугодия ближайшего следующего, т.е. пятого, года, рассчитывается как сумма объемов потребления

электроэнергии в I и во II кварталах пятого года, соответственно F_{17} и F_{18}. Для определения трендовой компоненты воспользуемся уравнением тренда

$$T = 5{,}7155 + 0.1864 \cdot t.$$

Получим:

$$T_{17} = 5{,}7155 + 0.1864 \cdot 17 = 8{,}8843;$$

$$T_{18} = 5{,}7155 + 0.1864 \cdot 18 = 9{,}0707.$$

Значения сезонной компоненты равны: $S_1 = 0{,}581$ (I квартал); $S2 = -1{,}977$ (II

квартал).

Таким образом,

$$F_{17} = T_{17} + S_1 = 8{,}8843 + 0{,}581 = 9{,}4653;$$

$$F_{18} = T_{18} + S_2 = 9{,}0707 - 1{,}977 = 7{,}0937.$$

Прогноз объема потребления электроэнергии на первое полугодие ближайшего следующего (пятого) года составит:

$$(9{,}4653+7{,}0937)=16{,}559 \text{ млн кВт/ч.}$$

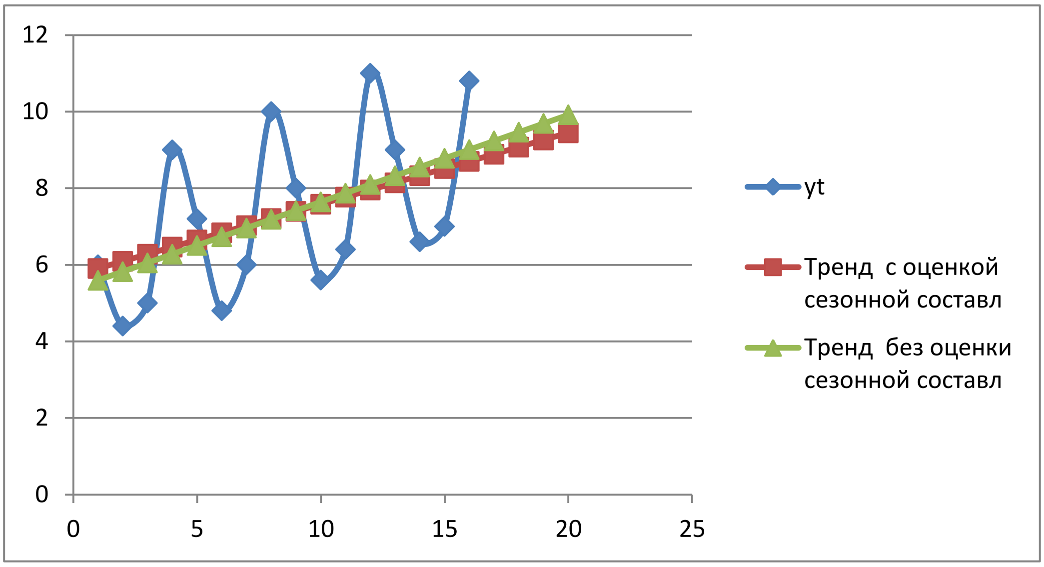


Рис 4.5 Разница трендов с учётом и без учёта сезонной составляющей составляет 0,5 млн кВт/ч.

Контрольные вопросы

1. Методы анализа структуры динамических рядов.
2. Сущность метода скользящей средней.
3. Общий вид аддитивной и мультипликативной моделей, процесс их построения.
4. Каким образом осуществляется прогнозирование по аддитивной модели?

ПРАКТИЧЕСКОЕ ЗАНЯТИЕ № 5
МОДЕЛИРОВАНИЕ ОДНОМЕРНЫХ ВРЕМЕННЫХ РЯДОВ. ПОСТРОЕНИЕ МУЛЬТИПЛИКАТИВНОЙ МОДЕЛИ ПОТРЕБЛЕНИЯ ЭЛЕКТРОЭНЕРГИИ

Краткие сведения из теории

Моделирование сезонных и циклических колебаний

Существует несколько подходов к анализу структуры динамических рядов, содержащих сезонные или циклические колебания. Поскольку моделирование циклических колебаний в целом осуществляется аналогично моделированию сезонных колебаний, мы рассмотрим только методы моделирования последних.

Простейший подход – расчет значений сезонной компоненты ***методом скользящей средней*** и построение аддитивной или мультипликативной модели временного ряда.

Общий вид ***мультипликативной модели*** выглядит так:

$Y = T \cdot S \cdot E$.

Эта модель предполагает, что каждый уровень временного ряда может быть представлен как произведение трендовой (T), сезонной (S) и случайной (E) компонент.

Выбор одной из двух моделей осуществляется на основе анализа структуры сезонных колебаний. Если амплитуда колебаний приблизительно постоянна, строят аддитивную модель временного ряда, в которой значения сезонной компоненты предполагаются постоянными для различных циклов. Если амплитуда сезонных колебаний возрастает или уменьшается, строят мультипликативную модель временного ряда, которая ставит уровни ряда в зависимость от значений сезонной компоненты.

Построение аддитивной и мультипликативной моделей сводится к расчету значений T, S и E для каждого уровня ряда.

Процесс построения модели включает в себя следующие шаги:

Шаг 1. Выравнивание исходного ряда методом скользящей средней.

Шаг 2. Расчет значений сезонной компоненты S.

Шаг 3. Устранение сезонной компоненты из исходных уровней ряда и получение выравненных данных ($T + E$) в аддитивной или ($T \cdot E$) в мультипликативной модели.

Шаг 4. Аналитическое выравнивание уровней ($T + E$) или ($T \cdot E$) и расчет значений T с использованием полученного уравнения тренда.

Шаг 5. Расчет полученных по модели значений ($T + E$) или ($T \cdot E$).

Шаг 6. Расчет абсолютных и/или относительных ошибок.

Если полученные значения ошибок не содержат автокорреляции, ими можно заменить исходные уровни ряда и в дальнейшем использовать временной ряд ошибок E для анализа взаимосвязи исходного ряда и других временных рядов.

Подробнее методику построения каждой из моделей рассмотрим на примерах.

Таблица 4.13

Год \ Квартал	I	II	III	IV
1	72	100	90	64
2	70	92	80	58
3	62	80	68	48
4	52	60	50	30

Рис. 4.5. Прибыль компании

Пример 1. Построение мультипликативной модели временного ряда.

Пусть имеются поквартальные данные о прибыли компании за последние четыре года (табл. 4.13).

График данного временного ряда (рис. 4.5) свидетельствует о наличии сезонных колебаний (период колебаний равен 4) и общей убывающей тенденции уровней ряда. Прибыль компании в весенне-летний период выше, чем в осенне-зимний период. Поскольку ампли-туда сезонных колебаний уменьшается, можно предположить существование мультипликативной модели. Определим ее компоненты.

Шаг 1. Проведем выравнивание исходного ряда методом скользящей средней. Методика, применяемая на этом шаге, полностью совпадает с методикой аддитивной модели. Результаты расчетов оценок сезон-ной компоненты представлены в табл. 4.14.

Шаг 2. Найдем оценки сезонной компоненты как частное от деления фактических уровней ряда на центрированные скользящие средние (гр. 6 табл. 4.14):

$$90 : 81{,}25 = 1{,}108;$$

$$64 : 80 = 0{,}80 \text{ и т.д.}$$

Используем эти оценки для расчета значений сезонной компоненты S (табл. 4.15). Для этого найдем средние за каждый квартал оценки сезонной компоненты S_t. *Взаимопогашаемость сезонных воздействий в* *мультипликативной модели выражается в том, что сумма значений сезонной компоненты по всем кварталам должна быть равна числу периодов в цикле.* В нашем случае число периодов одного цикла (год) равно четырем (4 квартала).

Год	№ квартала, t	Прибыль компании, y_t	Итого за четыре квартала	Скользящая средняя за четыре квартала	Центрированная скользящая средняя	Оценка сезонной компоненты
	1	2	3	4	5	6
1	1	72	---	---	---	---
	2	100	326	81,5	---	---
	3	90	324	81,0	81,25	1,108
	4	64	316	79,0	80,00	0,800
2	5	70	306	76,5	77,75	0,900
	6	92	300	75,0	75,75	1,215
	7	80	292	73,0	74,00	1,081
	8	58	280	70,0	71,50	0,811
3	9	62	268	67,0	68,50	0,905
	10	80	258	64,5	65,75	1,217
	11	68	248	62,0	63,25	1,075
	12	48	228	57,0	59,50	0,807
4	13	52	210	52,5	54,75	0,950
	14	60	192	48,0	50,25	1,194
	15	50	---	---	---	---
	16	30			---	---

Имеем:

$$0{,}918 + 1{,}209 + 1{,}088 + 0{,}808 = 4{,}023\ .$$

Определим корректирующий коэффициент:

$$k = \frac{4}{4{,}023} = 0{,}9943\ .$$

Определим скорректированные значения сезонной компоненты, умножив ее средние оценки на корректирующий коэффициент k.

$$S_i = \bar{S}_i \cdot k, \quad i = \overline{1,4}\ .$$

Показатели	Год	№ квартала, i			
		I	II	III	IV
	1 2 3 4	--- 0,900 0,905 0,950	--- 1,215 1,217 1,194	1,108 1,081 1,075 ---	0,800 0,817 0,807 ---
Итого за i-ый квартал (за все годы)		2,755	3,626	3,264	2,424
Средняя оценка сезонной компоненты для i-го квартала, $\overline{S}_i$		0,918	1,209	1,088	0,808
Скорректированная сезонная компонента, S_i		0,913	1,202	1,082	0,803

Проверим условие «равенства четырем» суммы значений сезонной компоненты:

$0{,}913 + 1{,}202 + 1{,}082 + 0{,}803 = 4$.

Получим следующие значения сезонной компоненты:

I квартал: S_1 = 0,913;

II квартал: S_2 = 1,202;

III квартал: S_3 = 1,082;

IV квартал: S_4 = 0,803.

Занесем полученные значения скорректированной сезонной компоненты в табл. 4.15 для соответствующих кварталов каждого года (стр. 3).

Шаг 3. Разделим каждый уровень исходного ряда на соответствующие значения сезонной компоненты. Тем самым мы получим величины $T \cdot E = Y : S$ (гр. 4 табл. 4.16), которые содержат только тенденцию и случайную компоненту.

Год	t	y_t	S_i	$T \cdot E = y_t : S_t$	T	$T \cdot S$	$E = y_t : (T \cdot S)$	$E' = y_t - (T \cdot S)$	$(E')^2$
	1	2	3	4	5	6	7	8	9
1	1	72	0,913	78,8609	87,8036	80,1647	0,8982	-8,1647	66,6621
	2	100	1,202	83,1947	85,0285	102,2042	0,9784	-2,2042	4,8585
	3	90	1,082	83,1793	82,2533	88,9981	1,0113	1,0019	1,0038
	4	64	0,803	79,7011	79,4782	63,8210	1,0028	0,1790	0,0321
2	5	70	0,913	76,6703	76,7030	70,0299	0,9996	-0,0299	0,0009
	6	92	1,202	76,5391	73,9279	88,8613	1,0353	3,1387	9,8513
	7	80	1,082	73,9372	71,1527	76,9873	1,0391	3,0127	9,0766
	8	58	0,803	72,2291	68,3776	54,9072	1,0563	3,0928	9,5654
3	9	62	0,913	67,9080	65,6025	59,8950	1,0351	2,1050	4,4309
	10	80	1,202	66,5557	62,8273	75,5184	1,0593	4,4816	20,0845
	11	68	1,082	62,8466	60,0522	64,9764	1,0465	3,0236	9,1419
	12	48	0,803	59,7758	57,2770	45,9934	1,0436	2,0066	4,0262
4	13	52	0,913	56,9551	54,5019	49,7602	1,0450	2,2398	5,0166
	14	60	1,202	49,9168	51,7267	62,1755	0,9650	-2,1755	4,7330
	15	50	1,082	46,2107	48,9516	52,9656	0,9440	-2,9656	8,7949
	16	30	0,803	37,3599	46,1764	37,0797	0,8091	-7,0797	50,1220

Шаг 4. Определим компоненту T в мультипликативной модели. Для этого рассчитаем параметры линейного тренда, используя уровни ($T \cdot E$). Результаты аналитического выравнивания этого ряда представлены ниже:

Константа (a)	90,5787
Коэффициент регрессии (b)	-2,7751
Стандартная ошибка коэффициента регрессии	0,2252
R-квадрат	0,9156
Число наблюдений	16
Число степеней свободы	14

Уравнение тренда имеет следующий вид:

$$T = 90{,}5787 - 2{,}7751 \cdot t.$$

Подставляя в это уравнение значения $t = 1, \ldots, 16$, найдем уровни T для каждого момента времени (гр. 5 табл. 4.16). График уравнения тренда приведен на рис. 4.6.

Шаг 5. Найдем уровни ряда по мультипликативной модели, умножив уровни T на значения сезонной компоненты для соответствующих кварталов. Графически значения ($T \cdot S$) представлены на рис. 4.6.

Шаг 6. Расчет ошибки в мультипликативной модели производится по формуле

$$E = Y : (T \cdot S).$$

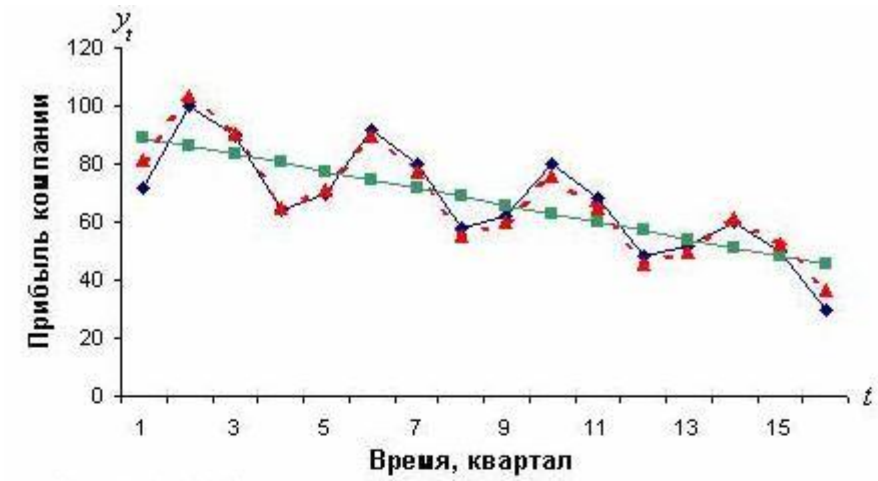


Рис. 4.6. Прибыль компании (фактические и выравненные по мультипликативной модели значения уровней ряда):
фактические значения;
тренд (T);
значения $(T \cdot S)$.

Численные значения ошибки приведены в гр. 7 табл. 4.16. Если временной ряд ошибок не содержит автокорреляции, его можно использовать вместо исходного ряда для изучения его взаимосвязи с другими временными рядами. Для того чтобы сравнить мультипликативную модель и другие модели временного ряда, можно по аналогии с аддитивной моделью использовать сумму квадратов абсолютных ошибок. Абсолютные ошибки в мультипликативной модели определяются как

$$E = Y - (T \cdot S).$$

В данной модели сумма квадратов абсолютных ошибок составляет 207,4007. Общая сумма квадратов отклонений фактических уровней этого ряда от среднего значения равна 5023. Таким образом, доля объясненной дисперсии уровней ряда равна: $(1 - 207{,}4007/5023) \cdot 100 = 95{,}871$, или 95,9%.

Выявление и устранение сезонного эффекта используются в двух направлениях. *Во-первых*, воздействие сезонных колебаний следует устранять на этапе предварительной обработки исходных данных при изучении взаимосвязи нескольких временных рядов. *Во-вторых*, это анализ структуры одномерных рядов динамики с целью прогнозирования уровней ряда в будущие моменты времени.

Пример 2. *Прогнозирование по мультипликативной модели*.

Предположим, по данным примера 4.5 необходимо сделать прогноз ожидаемой прибыли компании за первое полугодие ближайшего следующего года.

Прогнозное значение F_i уровня динамического ряда в мультипликативной модели в соответствии с соотношением $Y = T \cdot S \cdot E$ есть произведение трендовой и сезонной компонент. Для определения трендовой компоненты за каждый квартал воспользуемся уравнением тренда

$$T = 90{,}5787 - 2{,}7751 \cdot t.$$

Получим:

$$T_{17} = 90{,}5787 - 2{,}7751 \cdot 17 = 43{,}4013;$$

$$T_{18} = 90{,}5787 - 2{,}7751 \cdot 18 = 40{,}8282.$$

Значения сезонной компоненты равны: $S_1 = 0{,}913$ (I квартал); $S_2 = 1{,}202$ (II квартал).

Таким образом,

$$F_{17} = T_{17} \cdot S_1 = 43{,}4013 \cdot 0{,}913 = 39{,}6254;$$

$$F_{18} = T_{18} \cdot S_2 = 40{,}6262 \cdot 1{,}202 = 48{,}8326.$$

Прогноз ожидаемой прибыли компании на первое полугодие ближайшего следующего года составит:

$(39{,}6254 + 48{,}8326) = 88{,}458$ тыс.долл.

Контрольные вопросы

5. Методы анализа структуры динамических рядов.
6. Сущность метода скользящей средней.
7. Общий вид аддитивной и мультипликативной моделей, процесс их построения.
8. Каким образом осуществляется прогнозирование по мультипликативной модели?

ПРАКТИЧЕСКОЕ ЗАНЯТИЕ № 6
«ПОСТРОЕНИЕ МОДЕЛЕЙ АВТОРЕГРЕССИИ (AR)

Разберем пример прогнозирования курса акций Лукойла с помощью авторегрессионной модели. Эта модель так же называется AR (AutoRegressive). Эта модель используется во многих финансовых отраслях, где необходимо прогнозировать различные данные, например, прогнозирование значений ВВП, объема продаж товаров на предприятии, стоимости ценных бумаг и т.д. AR относится к классу регрессионных методов. Рассмотрим авторегрессию первого порядка AR(1), которая характеризует тесноту связи между соседними значениями ценового или иного ряда.

Авторегрессионная модель первого порядка имеет следующую формулу:

$$Y_i=\alpha+\beta*Y_{i-1}+\varepsilon,$$

Где β, α – коэффициенты авторегрессии;

ε– белый шум, независимая случайная величина;

Y_{i-1} – предыдущее значение временного ряда;

Y_i – текущее значения временного ряда.

Для того что бы сделать прогноз на основе этой модели воспользуемся программой MS Excel. И так, давайте спрогнозируем стоимость акций Лукойла (LKOH) на несколько периодов вперед. Построение авторегрессии имеет схожий алгоритм с автокорреляцией. Дневные котировки акции взяты за один месяц с 31 июля 2010г. по 31 августа 2010г. Следует заметить, что взята ценная бумага, торгуемая на бирже ММВБ.

Для начала экспортируем котировки с сайта finam.ru за выбранный период. Всего получилось 21 значение котировки. Экспорт в Excel будет выглядеть следующим образом.

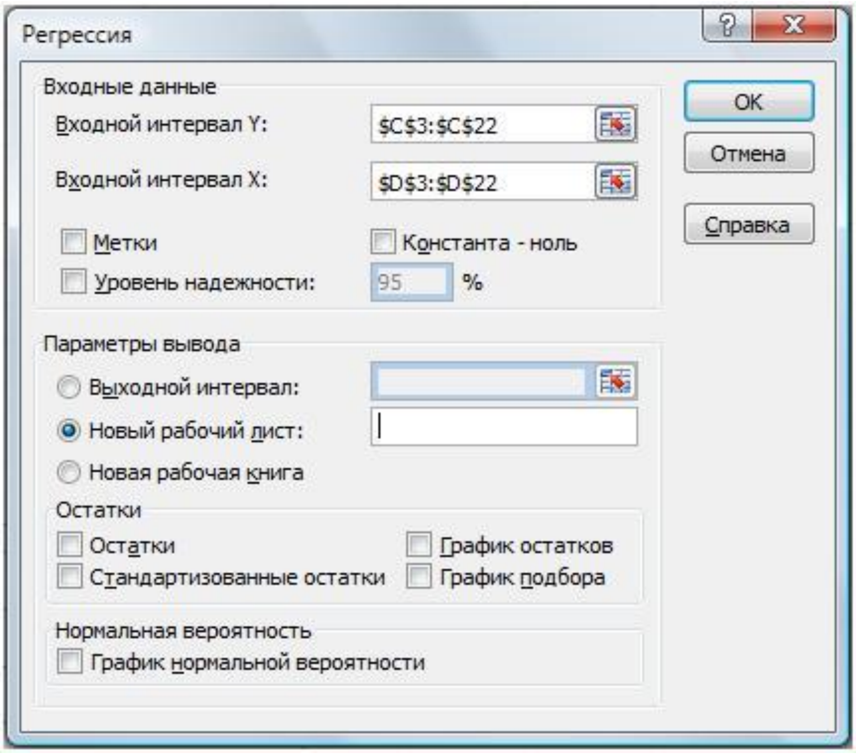
	A	B	C
1	Название	Дата	Close
2	LKOH	30.07.2010	1720
3	LKOH	02.08.2010	1724.92
4	LKOH	03.08.2010	1693.5
5	LKOH	04.08.2010	1684.47
6	LKOH	05.08.2010	1700

График ценной бумаги представляет собой ярко выраженный **линейный тренд**.

Сделаем прогноз этой ценной бумаги на следующие три периода, то есть на 3 дня вперед. Для этого необходимо найти авторегрессию нашего ценового ряда, то есть тесноту связи между соседними членами ценового ряда. Скопируем со сдвигом в одну ячейку наш временной ряд и вставим его в столбец “D”.

	A	B	C	D
1	**Название**	**Дата**	**Close**	
2	LKOH	30.07.2010	1720	
3	LKOH	02.08.2010	1724.92	1720
4	LKOH	03.08.2010	1693.5	1724.92
5	LKOH	04.08.2010	1684.47	1693.5

Далее, рассчитаем коэффициенты авторегрессии для ценового ряда Лукойла. Для расчетов коэффициентов воспользуемся надстройкой «Анализ Данных» и разделом «Регрессия». В поле «входной интервал Y» введем значения котировок из столбца “C”. В поле «входной интервал X» введем значения тех же котировок сдвинутых на один интервал. Следует заметить, что последнее значение у сдвинутого интервала “C23” и первое значение “C2” не входит в выделение.



После этого выйдет отчет по регрессии. Разберем более подробно этот отчет. Коэффициент *R- квадрат* показывает качество модели, чем выше это значение, тем лучше. *P-Значение* меньше 15%, значит коэффициенты AR(1), считаются значимыми. *Значимость F* равна 0 –это говорит о хорошем качестве всего уравнения. *Коэффициенты* это коэффициенты альфа (α) и бета (β) подобранные для нашей регрессионной модели тренда.

	A	B	C	D	E	F
1	ВЫВОД ИТОГОВ					
2						
3	*Регрессионная статистика*					
4	Множественный R	0.87911462				
5	R-квадрат	0.772842514				
6	Нормированный R-квадрат	0.760222654				
7	Стандартная ошибка	12.9765925				
8	Наблюдения	20				
9						
10	Дисперсионный анализ					
11		*df*	*SS*	*MS*	*F*	*Значимость F*
12	Регрессия	1	10312.35347	10312.35347	61.24017973	3.34956E-07
13	Остаток	18	3031.055154	168.391953		
14	Итого	19	13343.40862			
15						
16		*Коэффициенты*	*Стандартная ошибка*	*t-статистика*	*P-Значение*	*Нижние 95%*
17	Y-пересечение	332.45337	170.4003768	1.951013115	0.066798273	-25.54453673
18	Переменная X 1	0.799170454	0.102122442	7.825610502	3.34956E-07	0.584619166

Модель динамики ценной бумаги описывается следующим уравнением:

$$Y=332.45+0.79*Yt-1$$

Теперь построим непосредственно сам прогноз по этой модели. Для этого в колонке “E” введем формулу нашей авторегрессии AR(1). =332.45+0.79*D3

Авторегрессия будет строится только до 23 строчки, пока есть значения курса Лукойла. Далее необходимо прогнозировать уже от предыдущего прогноза, поэтому в ячейке “D24” введем формулу, берущую значения предыдущего прогноза:

“D24” =E23

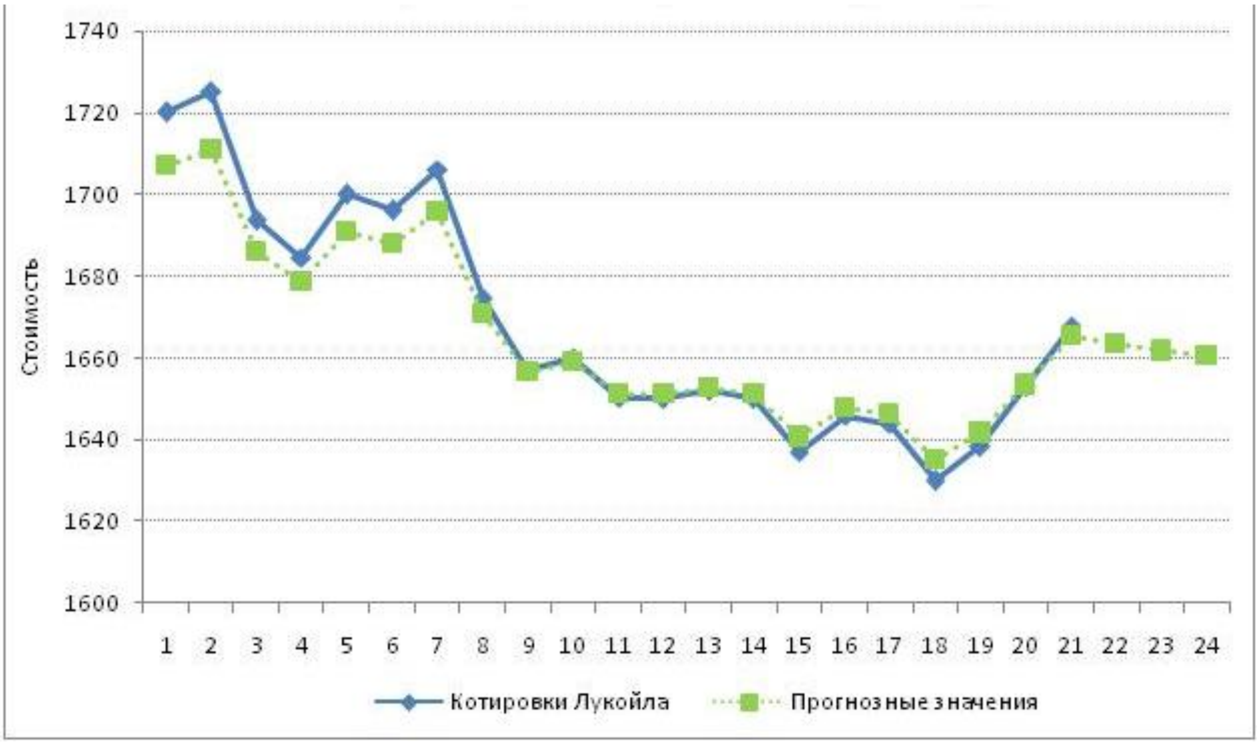
“D25” =E24

и т.д.

После полученных значений считаем значение нашей прогнозной модели для ячеек “E24 – E27”.

	A	B	C	D	E
1	Название	Дата	Close	Исх.ряд	Прогноз.
2	LKOH	30.07.2010	1720		
3	LKOH	02.08.2010	1724.92	1720	1707.027
4	LKOH	03.08.2010	1693.5	1724.92	1710.958
5	LKOH	04.08.2010	1684.47	1693.5	1685.849
6	LKOH	05.08.2010	1700	1684.47	1678.632
7	LKOH	06.08.2010	1696	1700	1691.043
8	LKOH	09.08.2010	1706.05	1696	1687.846
9	LKOH	10.08.2010	1674.67	1706.05	1695.878
10	LKOH	11.08.2010	1657	1674.67	1670.8
11	LKOH	12.08.2010	1660	1657	1656.679
12	LKOH	13.08.2010	1650.09	1660	1659.076
13	LKOH	16.08.2010	1650	1650.09	1651.157
14	LKOH	17.08.2010	1652.01	1650	1651.085
15	LKOH	18.08.2010	1650	1652.01	1652.691
16	LKOH	19.08.2010	1637.07	1650	1651.085
17	LKOH	20.08.2010	1645.87	1637.07	1640.751
18	LKOH	23.08.2010	1644	1645.87	1647.784
19	LKOH	24.08.2010	1630	1644	1646.29
20	LKOH	25.08.2010	1638.3	1630	1635.101
21	LKOH	26.08.2010	1652.99	1638.3	1641.734
22	LKOH	27.08.2010	1668	1652.99	1653.474
23				1668	1665.47
24				1665.47	1663.448
25				1663.448	1661.831
26				1661.831	1660.54
27				1660.54	

Построим значения исходного ряда и прогнозные значения на основе авторегрессии. Получится следующий график прогнозных значений (зеленый график).



Вывод

Использование регрессионных моделей позволяет построить довольно четкие прогнозные модели. В данной модели использовалась только линейная регрессия, для описания линейных трендов. Так же можно описывать движения ценных бумаг: экспоненциальными, логарифмическими или полиноминальными трендами. На рынках где присутствует сильная нелинейность и хаотичность, использование таких методов не принесет желаемого. Для прогнозирования нелинейных зависимостей используют различные нейронные сети.

Задание для самостоятельного решения по группам

Задание. На основании данных таблицы для соответствующего варианта построить модель авторегрессии и оценить ее качества.

1. Построить уравнение авторегрессии.
2. Проверить значимость уравнения регрессии.
3. Дать интерпретацию полученным значениям.
4. Проверить наличие автокорреляции в остатках.

Вариант	Номер графы табл. П1.3 для результативной переменной y	Номер графы табл. П1.3 для факторной переменной x	Уровень значимости
1	3	4	0,05
2	3	5	0,01
3	3	11	0,05
4	3	15	0,01
5	10	4	0,05
6	10	5	0,01
7	10	11	0,05
8	10	15	0,01
9	16	4	0,05
10	16	5	0,01
11	16	11	0,05
12	16	15	0,01
13	7	4	0,05
14	7	5	0,01
15	7	11	0,05
16	7	15	0,01
17	14	4	0,05
18	14	5	0,01
19	14	11	0,05
20	14	15	0,01
21	6	4	0,05
22	6	5	0,01
23	6	11	0,05
24	6	15	0,01
25	11	8	0,05

Пример выполнения практического занятия.

Исходные данные даны в таблице выше. Уровень значимости α=0,05

Данные наблюдений

Год наблюдения	Y_t	X_t	Y_{t-1}	Расчетные значения $\hat{Y}_{t-1}$
1	1016,6	1412,7	–	–
2	1435,9	1978,9	1016,6	1060,6
3	1776,1	2292,0	1435,9	1443,8
4	2003,8	2514,4	1776,1	1655,8

1	2	3	4	5
5	3265,7	4632,0	2003,8	1806,3
6	4476,9	7116,6	3265,7	3239,7
7	5886,9	8819,9	4476,9	4921,5
8	7443,2	10627,5	5886,9	6074,5
9	9024,8	12886,1	7443,2	7298,0
10	11401,4	16679,9	9024,8	8826,9
11	14363,5	21079,5	11401,4	11394,9
12	17742,6	26009,7	14363,5	14372,9

1) Построение уравнения авторегрессии.

$$y_t = a + b_0 \cdot x_t + c_1 \cdot y_{t-1} + \varepsilon_t .$$

Для введения инструментальной переменной построим уравнение регрессии

$$\hat{y}_{t-1} = d_0 + d_1 \cdot x_{t-1} + u_t ,$$

используя функцию «Сервис.Анализ данных.Регрессия» табличного процессора MS Excel (рис 3.2). Задавая соответствующие диапазоны данных в окне определения параметров регрессии, получим

Результаты регрессионного анализа

Показатели	Коэффициенты уравнения регрессии	Стандартная ошибка определения коэффициентов	t-статистика	Вероятность ошибки α	Нижние 95%–пределы	Верхние 95%–пределы
Y-пересечение	104,31	97,732	1,067	0,314	–116,778	325,394
Переменная X_{t-1}	0,677	0,009	71,398	0,000	0,655	0,698

Уравнение, определяющее инструментальную переменную $\hat{y}_{t-1}$ имеет вид

$$\hat{y}_{t-1} = 104{,}31 + 0{,}677 \cdot x_{t-1} . \tag{5.16}$$

Расчетные значения инструментальной переменной $\hat{y}_{t-1}$ приведены в таблице 5.2.

Используя функцию «Сервис.Анализ данных.Регрессия» табличного процессора MS Excel получим

множественный коэффициент корреляции $R = 0{,}9962$,

коэффициент детерминации $R^2 = 0{,}9993$,

$F_{факт} = 5737$,

уровень значимости уравнения регрессии $\alpha = 2{,}36 \cdot 10^{-13}$.

Результаты регрессионного анализа

Показатели	Коэффициенты уравнения регрессии	Стандартная ошибка определения коэффициентов	t-статистика	Вероятность ошибки α	Нижние 95%–пределы	Верхние 95%–пределы
Y-пересечение	139,80	82,870	1,687	0,130	-51,296	330,903
Переменная X_t	0,496	0,078	6,348	0,00022	0,316	0,676
Переменная $\hat{y}_{t-1}$	0,329	0,141	2,328	0,048	0,003	0,655

Уравнение регрессии

$$y_t = 139{,}80 + 0{,}496 \cdot x_t + 0{,}329 \cdot y_{t-1}. \qquad (5.17)$$

2) Проверка значимости.

Уравнение (5.17) значимо при $\alpha = 0{,}05$, так как его значимость $\alpha = 2{,}36 \cdot 10^{-13}$.

Из таблицы 4.4 следуют следующие значения уровней значимости значений параметров уравнения (4.13):

- параметр 139,80: $\alpha = 0, 130$;
- параметр 0,496: $\alpha = 0{,}00022$;
- параметр 0,329: $\alpha = 0{,}048$.

Следовательно при уровне значимости $\alpha = 0{,}05$ параметр 139,80 – не значим, а параметры 0,496 и 0,329 – значимы.

3) Интерпретация значений параметров уравнения.

Краткосрочный мультипликатор $b_0 = 0{,}496$.

Долгосрочный мультипликатор $b = \dfrac{b_0}{1 - c_1} = \dfrac{0{,}496}{1 - 0{,}329} = 0{,}739$.

Таким образом, увеличение x_t на 1 единицу приводит к росту y_t в том же периоде в среднем на 0,496 единиц. Долгосрочное изменение y_t составит 0,739 единиц, т. е. изменение x_t на 1 единицу в каком-либо периоде приведет к изменению y_t в долгосрочной перспективе в среднем на 0,739 единиц.

4) Проверим наличие автокорреляции в остатках для уравнения (5.17) с помощью критерия h Дарбина.

Результаты расчетов значений инструментальной переменной $\hat{y}_{t-1}$ по уравнению (5.16) и остатки показаны в таблице 5.5.

Вычислим величину d (5.14)

$$d = \frac{\sum_{i=2}^{n}(e_t - e_{t-1})^2}{\sum_{i=1}^{n} e_t^{\,2}} = \frac{363417}{202262{,}6} = 1{,}762.$$

Квадрат стандартной ошибки коэффициента при переменой y_{t-1} в (5.17)

$$V = 0{,}141^2 = 0{,}02.$$

Результаты расчетов для проверки автокорреляции остатков

Год наблюдения	Порядковый номер Переменная t	Данные наблюдений Y_t	Расчетные значения $\hat{Y}_t$	Остатки $e_t = \hat{Y}_t - Y_t$	e_t^2	$e_t - e_{t-1}$	$(e_t - e_{t-1})^2$
1		1016,6	–	–	–	–	–
2	1	1435,9	1470,2	34,3	1173,6	–	–
3	2	1776,1	1751,6	-24,5	601,2	-58,8	3454,78
4	3	2003,8	1931,6	-72,2	5208,9	-47,7	2270,81
5	4	3265,7	3031,1	-234,6	55015,6	-162,4	26367,5
6	5	4476,9	4735,0	258,1	66608,5	492,6	242694

1	2	3	4	5	6	7	8
7	6	5886,9	6133,2	246,3	60677,1	-11,8	138,277
8	7	7443,2	7409,1	-34,1	1165,1	-280,5	78658,4
9	8	9024,8	8931,8	-93,0	8655,4	-58,9	3469,30
10	9	11401,4	11316,2	-85,2	7265,8	7,8	60,7634
11	10	14363,5	14343,1	-20,4	418,1	64,8	4197,92
12	11	17742,6	17768,0	25,4	646,8	45,9	2105,04
Сумма					206262,6		363417

Вычислим величину критерия h (5.13)

$$h = (1 - \frac{d}{2}) \cdot \sqrt{\frac{n}{1 - n \cdot V}} = (1 - \frac{1,762}{2}) \cdot \sqrt{\frac{12}{1 - 12 \cdot 0.02}} = 0,473.$$

Определим значения $t_{\alpha/2}$ и $t_{1-\alpha/2}$ квантилей порядка α/2 и1–α/2 стандартизованного нормального распределения при уровне значимости α = 0,05:

$t_{\alpha/2}$ = НОРМСТОБР(α/2) = НОРМСТОБР(0,05/2) = – 1,96;

$t_{1-\alpha/2}$ = НОРМСТОБР(1–α/2) = НОРМСТОБР(1–0,05/2) = 1,96.

Так как выполняется условие

$$t_{\alpha/2} = -1,96 < h = 0,473 < t_{1-\alpha/2} = 1,96,$$

то делаем вывод об отсутствии автокорреляции в остатках для уравнения (5.17).

Результаты

1) Уравнение авторегрессии

$$y_t = 139,80 + 0,496 \cdot x_t + 0,329 \cdot y_{t-1}.$$

2) Проверка значимости.

Уравнение (5.17) значимо при α = 0,05.

При уровне значимости α = 0,05 параметр 139,80 не значим, а параметры 0,496 и 0,329 – значимы.

3) Интерпретация значений параметров уравнения.

Краткосрочный мультипликатор b_0 = 0,496.

Долгосрочный мультипликатор b = 0,739.

4) Проверка наличия автокорреляции в остатках для уравнения (5.17).

Автокорреляции в остатках отсутствует.

Контрольные вопросы

1. Дать понятие авторегрессионной модели.
2. Области применения авторегрессионного анализа.